JAN 6 1922 GENERAL LIBRARY
JAN - 7 1922
UNIV. OF MICH.

Journal of Educational Research

OFFICIAL ORGAN OF THE NATIONAL ASSOCIATION OF DIRECTORS OF EDUCATIONAL RESEARCH

B. R. BUCKINGHAM, EDITOR

ASSOCIATE EDITORS

E. L. ASHBAUGH W. W. CHARTERS GEORGE D. STRAYER
GUY M. WHIPPLE S. A. COURTIS LEWIS M. TERMAN

VOL. IV DECEMBER, 1921 No. 5

CONTENTS

COMPARING THE EFFICIENCY OF SPECIAL TEACHING METHODS BY MEANS OF STANDARDIZED TESTS. . . . . S. S. Brooks 337
UNRELIABILITY OF INDIVIDUAL SCORES IN MENTAL MEASUREMENTS. . . . . John L. Stenquist 347
THE EFFECT OF KINAESTHETIC FACTORS IN THE DEVELOPMENT OF WORD RECOGNITION IN THE CASE OF NON-READERS. . . . . Grace M. Fernald and Helen Keller 355
AN EXPERIMENT CARRIED ON WITH THE PUPILS OF THE RUSSELL PREVOCATIONAL ROOM. . . . . J. H. Voorhees 378
THE RELIABILITY OF PREDICTION OF PROPORTIONS ON THE BASIS OF RANDOM SAMPLING. . . . . Ben D. Wood 390
A SERIES OF STANDARDIZED DIAGNOSTIC TESTS FOR THE FUNDAMENTALS OF ELEMENTARY ALGEBRA. . . . . H. R. Douglass 395
SCALE OF ATTAINMENT NO. 3—FOR MEASURING "ESSENTIAL ACHIEVEMENT" IN THE THIRD GRADE. . . . . Luella W. Pressey 404
AN ANALYSIS OF THE CONTENT OF SIX THIRD-GRADE ARITHMETICS. . . . . F. T. Spaulding 413
EDITORIALS. 424
REVIEWS AND ABSTRACTS. 427
NEWS ITEMS AND COMMUNICATIONS. 433
NATIONAL ASSOCIATION OF DIRECTORS OF EDUCATIONAL RESEARCH. 439

Entered as second-class matter at the Post Office at Menasha, Wisconsin, under the act of March 3, 1879.

Journal of Educational Research

Copyright 1921, by
Public School Publishing Co.

Official Organ of the National Association of Directors of Educational Research

Editor-in-Chief: **B. R. Buckingham**
Director of Bureau of Educational Research, Ohio State University, Columbus, Ohio

Associate Editors

E. J. ASHBAUGH, *Ex-Officio, Secretary of Research Association, Ohio State University.*
W. W. CHARTERS,
Carnegie Inst. Tech., Pittsburgh.
S. A. COURTIS,
City Schools, Detroit, Mich.
GEORGE D. STRAYER,
Columbia University, New York City.
LEWIS M. TERMAN,
Leland Stanford Junior University, Stanford University, California.
GUY M. WHIPPLE,
College of Education, University of Michigan, Ann Arbor, Michigan.

Contributing Editors

VIRGIL E. DICKSON,
Director, Bureau of Research and Guidance, Public Schools, Oakland and Berkeley, California.
CHARLES FORDYCE,
Director of Bureau of Educational Measurements, University of Nebr., Lincoln, Nebr.
M. E. HAGGERTY,
Dean of the College of Education, University of Minnesota, Minneapolis.
V. A. C. HENMON,
Director, School of Education, University of Wisconsin, Madison.
ARTHUR W. KALLOM,
Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston, Mass.
F. J. KELLY,
Dean of University Administration, University of Kansas, Lawrence, Kan.
ROBT. H. LANE,
Assistant Superintendent of Schools, Los Angeles, Cal.
WILLIAM A. McCALL,
Assistant Professor of Education, Teachers' College, Columbia University.
GEORGE W. MELCHER,
Director, Bureau of Research and Efficiency, Public Schools, Kansas City, Missouri.
EUGENE A. NIFENECKER,
Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City.
JOSEPH P. O'HERN,
Assistant Superintendent of Schools and Director of Efficiency Bureau, Public Schools, Rochester, New York.
LEON. O. SMITH,
Assistant Superintendent (in charge of Educational Research), Public Schools, Omaha, Nebr.
W. W. THEISEN,
Director, Division of Reference and Research, Public Schools, Cleveland, Ohio.
CLIFFORD WOODY,
Director Bureau of Tests and Measurements, University of Michigan, Ann Arbor, Mich.

ARTICLES TO APPEAR SOON

HINES, HARLAN C.: WHAT LOS ANGELES IS DOING WITH THE RESULTS OF TESTING.

MINER, J. B.: AN AID TO THE ANALYSIS OF VOCATIONAL INTERESTS.

MINOR, RUBY: PRINCIPLES OF METHOD APPLIED IN CONCRETE SITUATIONS.

ROGERS, BERTHA M.
and
BAKER, TERESA: A DIAGNOSTIC AND REMEDIAL ACTIVITY IN SUPERVISION.

BROOKS, SAMUEL S.: SOME USES FOR INTELLIGENCE TESTS.

JOURNAL *of* EDUCATIONAL RESEARCH

VOLUME IV DECEMBER, 1921 NUMBER 5

COMPARING THE EFFICIENCY OF SPECIAL TEACHING METHODS BY MEANS OF STANDARDIZED TESTS[1]

SAMUEL S. BROOKS
District Superintendent, Winchester, New Hampshire

In the last article five principal factors in a teacher's efficiency were distinguished—namely, (1) managing ability, (2) natural aptitude, (3) method of teaching, (4) interest and industry, and (5) personality. The position was taken that no one of these five factors can be accurately and objectively measured independently of any or all of the other factors.

Although *method* was one of the factors mentioned we nevertheless now propose to measure the efficiency of methods. Note, however, that we do not propose to do so independently of the other factors.

In general the efficiency of a teacher and the efficiency of her methods are pretty much inseparable. It is a mooted question whether or not there can be a good teacher without good teaching methods. We hear it argued, for example, that a good teacher with a poor method will accomplish more than a poor teacher with a good method. This argument implies that good teachers using poor methods may secure better results than poor teachers using good methods, in the same way that a good carpenter with few and poor tools can do a better job than can the novice with the best and most complete set of tools obtainable. We must admit that there is much truth in the argument. Sometimes we find that a teacher who is ignorant of approved methods but who has great natural ability is obtaining better results than another teacher who is without natural aptitude but who, perhaps with all the advantages of normal training, is using or rather misusing,

[1] This is the seventh article by Superintendent Brooks on the general topic, "Putting Standardized Tests to Practical Use in Rural Schools."

the most approved modern methods. One has the true teaching instinct and ability to apply general principles and the other lacks these advantages.

Whatever may be the actual relations between good and poor teachers and good and poor methods, we can all agree, I think, that the best teachers are those who combine natural aptitude with thorough knowledge of up-to-date methods together with skill in applying them so as to realize their possibilities. And although we cannot measure the efficiency of a teacher's methods entirely apart from consideration of her general ability, there is a way, nevertheless, by which we can with the help of standardized tests obtain fairly accurate comparisons of the efficiency of various special methods, taking at the same time full cognizance of the teacher's general ability.

This can be done somewhat as we solve simultaneous equations in algebra—that is, by manipulating the various quantities so as to eliminate all but one of the unknowns. The value of the remaining unknown is readily found after the others are equalized so as to cancel each other. Yet it cannot be said that the eliminated quantities are ignored. The manipulations required to bring about the conditions suitable for their elimination give them their full force in evaluating the result.

And so, if we are to find the relative values of two or more special teaching methods, we must equalize as far as possible the conditions under which those methods are tried out, thus eliminating all the unknown quantities but one. The chief of these external conditions that would affect the accuracy of our results are the general ability of the teachers, the average mental abilities of the several groups of pupils, and the time devoted to class work with the method.

Now, there are two ways in which we may want to compare methods. We may want to discover which of two or more special methods of teaching a subject will give the best results when used by teachers of equal general ability or we may want to learn which of two or more special methods can be used to best advantage by a certain teacher.

To illustrate the first case, suppose we wanted to compare the results of drill in the fundamental operations of arithmetic as conducted in the usual more or less unorganized manner and without much regard for the special difficulties involved in definite types of

examples, with results of drill in the same operations by means of the Courtis Standard Practice Tests. To do this we should first choose our teachers for the trial. Their general ability should be as nearly equal as possible in order to eliminate so far as may be any inaccuracy in our conclusions due to differences in ability. Two teachers with approximately equal ratings by the method described in my last article would serve admirably. One should have had no experience with, and if possible no knowledge of, the Courtis Standard Practice Tests or of similar practice material, while the other should have had experience in their use and knowledge of their basic principles. It would not do to have the same teacher try to handle both methods because, on the one hand, if she had had experience with the practice tests, the defects of the haphazard procedure would be largely nullified by her knowledge of the principles underlying them; while, on the other hand, if she did not have such knowledge and experience, the advantages of the Courtis method would in some measure be lost.

The next step is to choose two groups of pupils. These groups should be neither too large nor too small; neither large enough to be cumbersome to handle as a class nor small enough to make average scores meaningless. From ten or twelve to twenty in a group is probably about right. The pupils in both groups should be in the same grade, and the average mental ages and average intelligence quotients of the two groups should be as nearly equal as possible. The pupils' mental ages and intelligence quotients are obtained of course by means of intelligence tests, some uses of which will be discussed in a later article.

As soon as the pupils have been selected, they should be carefully tested by means of standardized tests in the fundamentals, and their scores should be recorded. The testing of both groups and the scoring of the papers should be done by the same person, preferably a person experienced in such work. The period of drill should begin as soon as the tests have been given. Care should be taken to see that, in each group, exactly the same amount of time is devoted to drill in the fundamental operations each day. At the end of eight or ten weeks the tests should be given again, the scores recorded, and the progress of the two groups compared. The difference in progress of the two groups will approximate the difference in efficiency of the two methods.

The degree of accuracy of the results will depend upon the care with which the tests are given and the degree to which the conditions of the drill work are equalized. It is an open question whether or not the teachers themselves should be informed of the main purpose in view—that is, the purpose of comparing the efficiency of the two methods. If we could be perfectly sure that both teachers would be thoroughly interested and honest about the experiment it would undoubtedly be wise to seek their intelligent cooperation, since by so doing we should be more likely to get the best possible results from both methods. But if, thinking that their reputations are at stake, one or both are likely to be tempted to stretch the time limit for daily drill or to persuade the pupils to drill themselves for speed and accuracy outside of class, then it will probably be better to leave them in blissful ignorance of the main plot, merely seeing to it that each teacher devotes the same amount of time to class drill in the fundamentals each day. In this way one can infer what each of the methods would accomplish under everyday working conditions in the hands of equally competent teachers. If one is particularly desirous of getting the best results of which either method is capable, this purpose may perhaps be accomplished by asking each teacher separately to do her very best.

This particular problem was worked out in my district last year with rather interesting and fairly conclusive results. The Courtis Standard Practice Tests were not in use in the district but, wishing to introduce them the following September, I planned ahead to have the stage set for their appearance. That is, before the practice tests were introduced generally, I wanted if possible to prove definitely that better results could be accomplished by their use with less drudgery for both teachers and pupils.

This was before the teachers' ratings had been computed so I did not have their ratings for guides in selecting the teachers to carry out the experiment. But I did have the records of progress for each school as shown by the September and February tests. Wishing to secure as accurate results as possible under the circumstances, I tried the experiment in each of three different towns. To handle the work with the practice tests three teachers were selected (one in each town) who had shown interest and capability in adapting new ideas to classroom use and whose schools had made normal progress during the first half of the year. Five

weeks before the end of the winter term these three teachers were furnished with the Courtis Standard Practice Tests, Teacher's Manuals, and Students' Practice Pads. I showed them how to use the tests, pointed out their advantages, and explained the principles underlying them. Then I told them that for special reasons of which they would be informed in due time, I was anxious to have them become as expert as possible in using the tests by the end of that term. They assured me that they would do their best and I believe they did. At any rate, they did exceedingly well.

The other three teachers, one in each town, were chosen because their schools had also shown about normal progress for the first half of the year, and because of the further fact that they were all teachers of many years' experience, somewhat set in their ways and not taking kindly to new ideas, but withal hardworking, trustworthy, and capable of doing very good work in their own ways. In other words, they were good old-fashioned teachers.

The intelligence tests had been given by this time throughout the district and I hastened to get the mental ages and intelligence quotients of all the pupils for use in selecting the several groups. They were finally selected according to the plan outlined above except that the grades in any one school were too small to allow of groups of ten pupils being selected from the same grade in such a way that the six groups would all average the same in both mental ages and I. Q.'s. However, the conditions regarding M. A.'s and I. Q.'s were strictly observed and allowed for. The lowest mental age in any group was 10 years, 9 months and the highest was 11 years, 5 months. The I.Q.'s ranged from 97 to 105.

Using the Woody scales for measuring the ability of pupils in the fundamental operations, I gave the first tests to the six picked groups during the first week of the spring term, and corrected and scored them myself, tabulating the average scores for each group in each subject as shown in Table I, in the columns marked "A." At the time I gave the tests to each group of pupils, I had a talk with their teacher, telling her that for very important reasons I wanted her to see how much improvement she could bring about in that particular group during the ensuing twelve weeks by drilling the pupils together just fifteen minutes each day for speed and accuracy in the fundamental operations of arithmetic. The

TABLE I. AVERAGE SCORES IN THE WOODY SCALES

(a) Groups not Using Practice Tests

Operation	Group 1		Group 2		Group 3		Averages	
	A	B	A	B	A	B	A	B
Addition	11.6	14.8	12.0	14.7	11.8	15.1	11.8	14.8
Subtraction	8.2	10.6	8.4	10.3	7.9	10.5	8.2	10.5
Multiplication	8.5	12.2	8.3	12.0	8.1	12.4	8.3	12.2
Division	5.5	8.5	5.9	8.0	6.1	8.8	5.8	8.4
Mixed fundamentals	21.0	25.9	21.0	26.1	22.2	27.0	21.4	26.3

(b) Groups Using Practice Tests

Operation	Group 4		Group 5		Group 6		Averages	
	A	B	A	B	A	B	A	B
Addition	11.9	16.0	11.7	15.8	11.7	16.3	11.8	16.0
Subtraction	8.1	12.4	8.6	12.2	8.0	11.9	8.2	12.2
Multiplication	8.0	15.3	9.0	15.5	8.4	14.8	8.4	15.2
Division	5.4	9.6	5.7	10.2	5.8	9.3	5.6	9.7
Mixed fundamentals	22.0	29.0	19.5	30.0	23.0	29.6	21.5	29.5

three teachers trained for the purpose were directed to use only the Courtis Standard Practice Tests for the drill, but to use them for all they were worth. None of the teachers had any inkling of the real object in view. Yet each one was keyed up to do her best after her own fashion. Every pupil in the six groups was promised a special holiday for not missing more than one day during the term. Pedagogically, of course, this may have been questionable, but psychologically it proved very effective; and I hoped that the end would justify the means. At any rate, I know that a large majority of the pupils got their holiday.

The work was supervised as closely as possible throughout the term. Neither from observation nor by questioning pupils could I detect any evidence that the rules of the game were being disregarded by any of the teachers. At the end of twelve weeks the pupils were again tested with the Woody scales. The average scores for each group were placed in the "B" columns of Table I in such a way that each group's second score in each subject was opposite its first score in the same subject. According to the

table, the average score of the pupils of Group 1 on the first test in addition (Column A) was 11.6. The score for the same group in the second addition test was 14.8 as shown in the first "B" column. The scores for the three groups which did not use the practice tests were averaged for both first and second tests; and they are recorded in the fourth "A" and "B" columns, while in like manner the general averages for the three groups which used the practice tests are recorded in the last two columns of the table.

It will be noted that, according to the "A" columns of the general averages, the two main groups, the first consisting of the three smaller groups in which the practice tests were not used, and the second, of the three groups in which they were used, started almost exactly even in the race as might be expected under the circumstances. The first score of both groups in addition was 11.8 and the first score in subtraction for each group was 8.2. The remaining first scores differed by but one or two tenths of a unit. But this correspondence is no longer apparent when the "B" columns of general averages are considered. The final scores of the group using the practice tests are seen to be considerably larger than those of the group not using them. The differences between the scores contained in the fourth and last "B" columns represent the difference in progress of the two main groups.

The group of pupils drilled with the practice tests has all the best of the argument, the difference in progress being sufficiently great to prove conclusively considerable superiority for the Courtis method properly handled. On the whole, the improvement of all the groups was surprisingly large for a period of only twelve weeks. It amounted on the average to about a year of progress for the groups which did not use the practice tests and to about a year and three-quarters for the group using the practice tests. This merely goes to show what can be accomplished by intensive work along definite lines when the interest of teachers and pupils has been thoroughly aroused.

Now to return to the second way in which we might want to compare special methods. Suppose we wish to learn which of two or three special methods will give the best results with a particular teacher. This is quite a different matter from measuring the relative efficiency of the methods themselves. Only in exceptional cases can methods be accurately compared when handled by the same teacher. For such a purpose the teacher must be equally skilled in the use of the methods to be compared

and without prejudice in favor of any particular method. In particular she must have a thorough knowledge of the special advantages and disadvantages of each method and know how to minimize the latter and make the most of the former. In no other way can the methods be given a fair trial. Only an exceptionally well-trained and widely-experienced teacher, with the impartial mind of a scientist seeking truth through experiment, can fulfill these conditions. Such teachers are not to be found in every school system.

We know that quite often a method of teaching which has proved highly successful when handled by its originator or by teachers specially trained by him, has failed miserably when introduced into an alien school system where the teachers were trained and experienced in other methods. And such failure is not to be wondered at. When the mere form of a new method, without its spirit, is introduced among workers lacking a knowledge of the proper technic to accompany the method, and naturally prejudiced in favor of their own methods, the new method is foredoomed to failure. A few of the better teachers, specially endowed with adaptability and initiative, may grasp the essential advantages of the new method, gradually evolve a suitable technic to fit it, and adopt it as their own. But most teachers, finding themselves accomplishing less with the new method than they did with the old, and longing for the familiar routine, will, unless constant supervision prevents, return surreptitiously at least to their former procedure, convinced that there is none better and that attempting new methods is a waste of time and trying to the nerves.

Of course if the real interest of the teachers can be aroused in the new method by a judicious advertising campaign before it is introduced, and if everybody's patience holds out long enough, and a definite policy of teacher-selecting and teacher-training is carried on, eventually the new method will come into its own if it really possesses marked advantages. But in too many instances the innovation is discarded as worthless after a few months of half-hearted trial without any adequate attempt having been made to modify the environment to fit the new method. And the chief factors contributing to such failures in attempting to introduce new methods of teaching into a school system are the indifference of teachers or their actual antagonism toward new methods, in general, their lack of knowledge concerning

particular new methods, and their lack of foresight and initiative in adapting themselves and their ideas to changing conditions. Probably the most annoying factor and the one most difficult to eliminate is the teacher's mental attitude toward new ways of doing things, her clinging to familiar trails, and her aversion to breaking new paths even in the interest of finding a smoother, shorter, and pleasanter road to her goal.

Hence new methods, unless real interest and belief in them have been aroused in the teacher beforehand, have to contend against ignorance and indifference or prejudice from the start. I repeat, therefore, that the efficiency of new methods cannot be accurately compared with that of old methods if the new ones are tested by the very teacher whose own methods are being questioned as to their comparative worth. Her attitude is too much like that of the hen defending her chickens from the hawk that would destroy them, the teacher's chickens being her own familiar methods while the hawk is the superintendent with his disturbing new ideas.

We can, however, determine pretty accurately which of two methods a teacher can (or will) handle most efficiently regardless of the actual possibilities inherent in the two methods. And since it is essential that each teacher use, in general, the methods with which she can produce the best results, it is also essential that we know what those methods are. It will not be found profitable, merely for the sake of having certain new methods, to enforce their continued use on teachers who cannot or will not produce as good results with them as they produce with their own methods. So we must have some way of determining whether or not teachers are doing as well or better with the new methods after using them a reasonable length of time, say six months or a year.

This can be done with the help of standardized tests. First select ten or a dozen pupils in the school with mental ages and I.Q.'s as nearly equal as it is possible to arrange. Divide them into two equal groups that average about the same in mental ability. Next test them with some standardized tests in the subject for which special methods are to be compared. Then have the teacher try out two methods, one on each group of pupils, over a period of three or four months. At the end of that time give the tests again and compare the progress of the two groups.

Such a trial will not necessarily prove which method is the better as regards possibilities, nor with which method the teacher *could* do better if she had the proper inclination and training, but it will prove with which method she *will* do better under existing conditions. And that is the essential point. If, after preparing carefully for the introduction of a new method of teaching some subject, by discussing its possibilities with teachers individually and collectively, and by furnishing them with suitable reading material concerning its basic principles, special advantages, and technic; if after demonstrating to the teachers the proper handling of the method and giving them a reasonable length of time to acquire skill in its use; and if after striving in every way to arouse their interest and hearty cooperation in giving the new method a thorough tryout; if after doing all these things and as many more as you can think of, you make such a comparison as outlined above and find that a teacher either cannot or will not do at least as good work with the new method as she did with the old, then it is time to discard either the teacher or the method. If your best teachers have succeeded in getting superior results by using the new method, it means that the method is all right and it may be wise to keep the method and get a new teacher. But if your best teachers have failed to get better results with the new method after several months of earnest effort, it will probably be better to discard the method.

At any rate, in order that the children may get the most for their time and the taxpayers the most for their money, it behooves us to make sure that the methods in use in the schools under our direction are the most efficient that can be used under existing circumstances. We can do this either by selecting and training teachers to fit our chosen methods or by selecting methods to fit the available teachers. Most emphatically it is not efficiency to cling to new methods forced upon untrained or improperly trained and often unwilling teachers just because they are up-to-date methods, when those teachers are not doing as good work with them as with their own methods. Unless we can train our teachers successfully in the proper use of the new methods, or obtain teachers already trained in their use, we had better stick to the old a little longer. Standardized tests will help to prove whether or not the new methods are more successful than the old methods in a particular environment. Results are more important than methods.

UNRELIABILITY OF INDIVIDUAL SCORES IN MENTAL MEASUREMENTS

JOHN L. STENQUIST
Bureau of Reference, Research, and Statistics, New York City Public Schools

With the rapid adoption of tests, including both measurements of intelligence and of educational achievement for purposes of better classification of pupils, a fundamental fallacy has in a large number of cases crept into the statistical procedure, the fallacy, namely, of neglecting to take account of the variability of individual scores.

Many users of tests have innocently passed from the former considerations of grade or school *averages* to the consideration of the status of *individual pupils* without making provision for the unreliability of a single measurement. Undoubtedly the status of the individual pupil is precisely what must be emphasized by school administrators in order to obtain any great benefit from the use of the tests. But caution must be exercised in taking this step lest we fall into serious blunders and reach unsound conclusions.

AVERAGE VS. INDIVIDUAL SCORES

When dealing with measurements of a group—say a class with an ordinary test of intelligence—we customarily compute the class average. The reliability of this average will generally be sufficiently high for all practical purposes. On the basis of it we are justified in inferring that Class A is brighter in general than Class B, that it is less bright than Class C, etc. If a second form of the same test is given, the average derived from it will not differ greatly from the average for the first form. Such class scores are fairly reliable because they are based upon a comparatively large number of measurements. In any such group measurement we really have as many samples of that group as there are individuals. The unreliability of each score is, for the most part, due to chance errors; and these discount each other in the long run and can, therefore, be neglected so far as the group is concerned.

When, however, we turn from the consideration of the score for the group to the consideration of the score of any individual pupil in the group, as determined by a single test, the trust-

worthiness of our score is entirely changed. In a second form of the same test, the score of a given pupil will not necessarily be closely similar to his score in the first form. This is due in part to lack of sufficient sampling of his ability. Many factors come in to change his score; thus he may be physically less fit during one test than during the other; he may be emotionally disturbed at one time and not at another; the precise tasks required in different forms of the same test may for accidental reasons be of unlike difficulty *for him*, even though they are identical, as shown by group averages. Hence we have the net effect of variations about an average performance. This phenomenon is probably as well known as any other in the whole field of mental measuring; yet, it is a common practice totally to disregard it. In attempting to classify individual pupils by test results, as is now being done in many cities, radical changes in grading are frequently recommended, entirely neglecting the unreliability due to insufficient sampling of each pupil's ability.

The Amount of Variation in Successive Tests

The actual variations which may and do occur in a pupil's score in successive tests of the same trait are a serious matter when we attempt to place him in a specific grade according to any one of such scores. To show concretely what happens, the results for a sample class taken at random out of some sixty or seventy classes in public school 64, Manhattan, is cited in Table I. The tests were given under closely parallel conditions during a period of about four weeks. They were all administered by the same teacher, who was one of the most skillful of forty or fifty who gave tests. All scoring was carefully verified by a recheck on every pupil. An examination of Table I shows the mental age as determined by each test, and the average age according to all five tests (Column 12). The average age is the most significant as it represents five distinct trials on the part of each pupil. Using this as the true mental age, the extent to which each pupil deviated from it in five trials is shown in columns 3, 5, 7, 9, and 11.

Table I reads: Pupil No. 1 obtained a mental age of 166 months according to the National Intelligence Test, Scale A. This was two months less than his mental age as obtained by averaging his results on all the tests he took; his mental age according to the National Intelligence Test, Scale B, was in excess of 180 months or (at least) twelve months more than his mental age as obtained by averaging his results on all the tests he took; etc.

TABLE I. RESULTS IN FIVE INTELLIGENCE TESTS GIVEN BY THE ONE EXAMINER TO ONE CLASS

Pupil No.	National Scale A	Deviation from Average	National Scale B	Deviation from Average	Otis Group Intelligence Scale	Deviation from Average	Haggerty Intelligence Delta 2	Deviation from Average	Thorndike Reading Scale	Deviation from Average	Average Age	Average Deviation
1	2	3	4	5	6	7	8	9	10	11	12	13
1	166	− 2	180+	+12	179	+11	146	−22	162	− 6	168	10 6
2	175	− 1	180+	+ 4	182	+ 6	180	+ 4	162	−14	176	5 8
3	180+	+ 9	180+	+ 9	158	−13	178	+ 7	159	−12	171	10
4	152	−10	162	0	162	0	163	+ 1	170	+ 8	162	3 8
5	160	− 5	169	+ 4	160	− 5	173	+ 8	162	− 3	165	5
6	161	− 4	180+	+15	156	− 9	164	− 1	163	− 2	165	6 2
7	181	+ 1	180+	0	176	− 4	180+	0	180	0	180	1 0
8	134	−10	159	+15	150	+ 6	132	−12	146	+ 2	144	9 0
9	151	−15			165	− 1	174	+ 8	175	+ 9	166	8 2
10	155	− 8	143	−20	170	+ 7	181	+18	174	+11	163	12 8
11	156	+13	150	+ 7	138	− 5			127	−16	143	10 2
12	140	−17	182	+25	152	− 5			155	− 2	157	12 2
13	160	+ 9	148	− 3	120	−31	155	+ 4	172	+21	151	13 6
14			147	+17	123	− 7	125	− 5	126	− 4	130	8 2
15	159	0	169	+10	152	− 7	172	+13	142	−17	159	9 4
16	160	− 6	180+	+14	163	− 3	170	+ 4	157	− 9	166	7 2
17	138	−25	180+	+17	176	+13	174	+11	134	−29	163	19 0
18	159	− 8	154	−13			177	+10	179	+12	167	10 6
19	180	− 4	180+	− 4	206	+22	180	− 4	175	− 9	184	8 6
20	150	− 6	160	+ 4			141	−15	172	+16	156	8 2
21			181	+ 4	178	+ 1	175	− 2	172	− 5	177	3 0
22			180+	+13	167	0	161	− 6	160	− 7	167	6 5
23	174	− 3	183	+ 6	180	+ 3	180	+ 3	166	−11	177	5 2
24	170	− 3	180+	+ 7			182	+ 9	180	+ 7	173	6 5
25	126	−30	185	+29	154	− 2	157	+ 1	157	+ 1	156	8 5
26	139	−13	161	+ 9	144	− 8	152	0	164	+12	152	8 4
27	140	−16	157	+ 1	186	+30	162	+ 6	137	−19	156	14 8
28	142	−18	180+	+20	146	−14	157	− 3	176	+16	160	14 2
29	147	−12	180+	+21	141	−18	173	+14	151	− 8	159	13 8
30	139	−25	180+	+16	182	+18	167	+ 3	154	−10	164	14 4
31	165	− 9	172	− 2			180+	+ 6	180+	+ 6	174	5 7
32	180	+ 5	180+	+ 5			175	0	166	− 9	175	4 7
33	147	+ 6	160	+19	132	− 9	124	−17	143	+ 2	141	10 6
Totals	4686	293	5462	345	4498	258	5110	217	5298	315	5367	295 9
Averages	156 2	9 76	170 6	10 7	160 6	9 2	164 8	7 0	160 5	9 6	162 6	9 0

Five Parallel Tests of General Intelligence

While other tests which are often utilized as intelligence measures were given, they are not here included because of their lower correlations with our criterion. This criterion is the composite (equally weighted) score in the following six tests:

1. The National Intelligence Test, Scale A.
2. The National Intelligence Test, Scale B.
3. The Haggerty Intelligence Test, Delta 2.
4. The Otis Intelligence Scale, Advanced.
5. Meyers Mental Measure.
6. Kelley-Trabue Language.

The coefficients of correlation for each of the tests in Table I with this criterion are:

National A r . . . . . $=0.801$ $(n=560)$
National B r . . . . . $=0.788$ $(n=518)$
Otis (Advanced) $r=0.680$ $(n=551)$
Haggerty, Delta 2 $r=0.808$ $(n=532)$
Visual Vocabulary $r=0.680$ $(n=461)$

In addition to our records in the tests given in Table I, we also have records in the Kelley-Trabue Language tests, in the Woody-McCall Arithmetic, and in Meyers Mental Measure. These, however, correlate with a coefficient of only 0.58 $(n=581)$, 0.39 $(n=298)$, and 0.48 $(n=544)$ respectively, with our criterion. National A, and B, and Haggerty are very similar in nature, and the Otis test does not differ greatly. Thorndike's Visual Vocabulary Test while purely a test of word knowledge, correlates exactly the same as the more elaborate Otis test. The spurious self-correlations do not appreciably operate to the disadvantage of any one test.

Equating Norms

In turning raw scores into mental ages (to secure a common unit) as has here been done, the question of the comparability of norms as furnished by the authors of the tests is involved. These were equated for all but the Visual Vocabulary Test[1] as follows: Twenty-seven classes including 1,007 pupils were measured by all of the tests; utilizing the published norms every raw score

[1] The norms for this test were established with great care and careful statistical procedure by Mr. R. H. Franzen, Director of Educational Research, Des Moines, Iowa, on the basis of some ten thousand cases in New York City and vicinity.

was then turned into a mental age. These twenty-seven classes included six eighth-grade, seven seventh-grade, ten sixth-grade, and four fifth-grade classes.

Considering these twenty-seven classes as a unit, each test should, if it measures the same thing as the others (as it does approximately), and if the norms are comparable, give the same mental age to this unit of 1,007 cases. This we did not find to be the case. Comparing the average age according to each of the tests with the average age according to the Haggerty Delta 2, we found that the average age according to the National A was 13 months higher, and that the corresponding figure for National B was 7.5 months higher, for the Otis 3.7 months higher, and for Myers Mental Measure 6.4 months lower.

The published norms in current usage at the time of our testing and the amount of deviation from Haggerty's Norms (which were taken as a point of reference, because of the large number of cases and their wide geographical distribution) are given in Table II.

When we say that the norms for the National Intelligence Test, Scale A, were thirteen months higher, we mean that the

TABLE II. PUBLISHED NORMS USED

Test	Age										Deviation of Norms from Haggerty Delta 2
	6	7	8	9	10	11	12	13	14	15	
National A			74	79	89	102	113	123	134		13 months higher
National B			76	81	89	102	112	124	128		7.5 months higher
Haggerty, Delta 2			25	43	55	66	77	87	100	115	
Otis Advanced			40	52	64	76	88	100			3.7 months higher
Meyers Mental Measure	9	17	23	28	34	39	43	47	49	49	6.4 months lower

critical score corresponding to each mental age, as given in the manual accompanying the test, was apparently too high, and that on the average it represented the children as being thirteen months younger mentally than they are represented to be by our revision. Accordingly, we lowered the norms for this test, by this amount thus making them less severe. In effect, this means that

for each child when tested by the National Intelligence Test, Scale A, we added thirteen months to the mental age which would have been obtained according to the author's norms. Similarly seven months were added to the ages of children for the National Intelligence Test, Scale B, and four months for the Otis Scale. In Table I these adjustments have been made.

Analysis of Variation in Five Tests by the Sample Class

With these facts in mind we may examine more in detail the scores of each pupil in Table I.

Column 13 shows the average deviation (in months of mental age) for each pupil. Twelve pupils out of thirty-three or about 35 percent deviated on the average in the five tests, ten months or more, which for practical purposes is about equivalent to one year. If we examine individual cases we find still more serious evidence of how precarious a procedure it is to place a pupil precisely by a single test. Note Pupil No. 12, for example. If his score in National A, had been our sole basis he would have been classified as forty-two months younger mentally than he is shown to be by National B, and seventeen months younger than the average of five tests. Again, note Pupil No. 28, who deviations in months of mental age are in order: −18, +20, −14, −3 and +16. It will make an enormous difference to this pupil whether he is classified by National A or by National B the difference in his mental age by these two tests being thirty-eight months, or over three years! Almost the same is true for pupils 17, 27, 29 and 30. Six extreme cases are also graphically shown in Figure 1. Here it is apparent at a glance that the variation is very great. A pupil by any one test may easily be misplaced by two years.

The Extreme Cases

The cases selected for Figure 1 are, to be sure, the most extreme ones. It is true that most of the cases deviate less. The average of the average deviations for the class is 8.7 months. But when we propose reclassification of pupils, we propose classifying *all of them*. All are tested, all are placed. But if as many as a third of them may be approximately a year older or younger mentally than a single test result shows; and if a few (here we have at least six out of thirty-five or some 15 to 20 percent), out of every class may be misplaced as much as three or more years, then our work is too unreliable to be utilized for classification purposes,

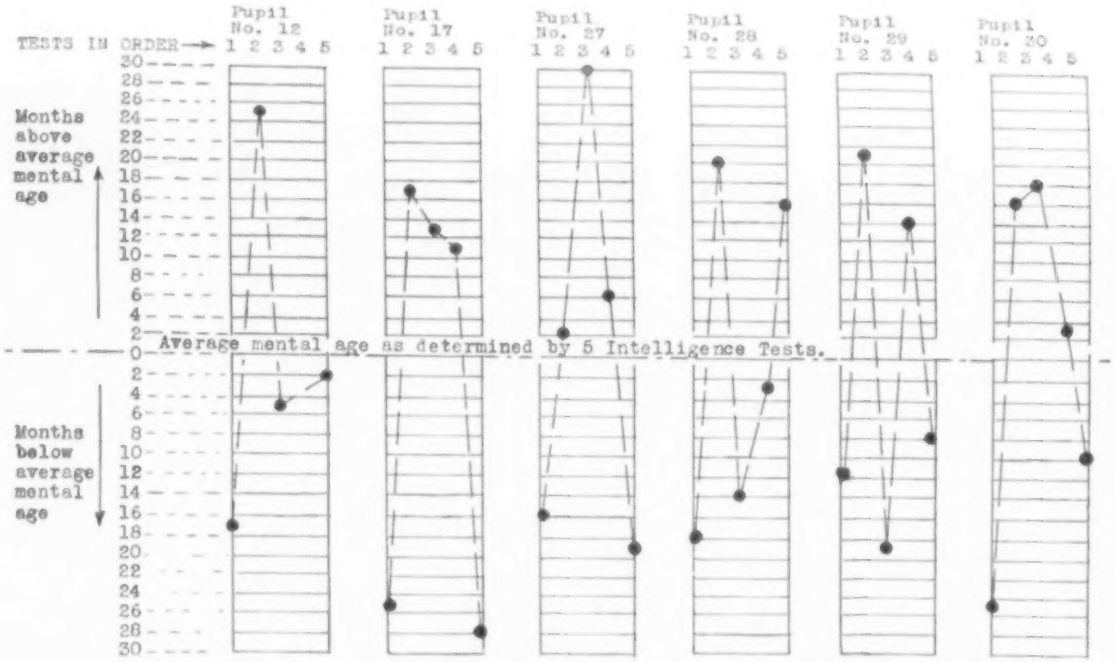
Pupil No. 12
Pupil No. 17
Pupil No. 27
Pupil No. 28
Pupil No. 29
Pupil No. 30
TESTS IN ORDER → 1 2 3 4 5
Months above average mental age
Months below average mental age
30 28 26 24 22 20 18 16 14 12 10 8 6 4 2 0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
Average mental age as determined by 5 Intelligence Tests.

FIGURE 1. SHOWING THE WIDE DIFFERENCES IN MENTAL AGE AS DETERMINED BY 5 SUCCESSIVE INTELLIGENCE TESTS. WHICH IS NEAREST THE TRUTH?

and may be no improvement over ordinary teacher-judgment methods. We believe that variability as great as that we have shown is not unusual. Those who doubt this should provide checks upon their results as we have done and first demonstrate that the great variability that we have found is the exception. It is here maintained that it is the rule; and that unless in utilizing tests to diagnose individuals, measures are taken to overcome it, the present testing movement will fall into disrepute, because tests misapplied cannot do what is claimed for them by enthusiasts.

In connection with this failure of successive test results to agree, we might appropriately raise the question: "Which score shall we accept as most significant?" We commonly assume that the average is nearest the truth. However, as Thorndike has pointed out on several occasions, in nearly all physical feats, we take not the average performance but the maximum. We are not interested, for instance, in the average height which a pole vaulter can clear, but the maximum height. Similarly it would not be unreasonable to use the maximum performance in mental feats. However, this is beside the main issue, namely, the variation in score from test to test.

The Remedy

The remedy is clear, and simple, though to be sure it is not one to elicit great enthusiasm. What is needed is a more thorough testing of each individual pupil—more tests, with equated norms, utilizing of course more and more reliable ones as they become available, and a more and more precise technic. When our results for an individual pupil are even approximately as reliable as our averages for classes then they will withstand any criticism and will be trustworthy. But this means more labor, more expert direction, more test material, and greater cost.

The very least that may be done is to supplement the single thirty-minute intelligence test, now so popular, by at least one other—better still by two or three—before extravagant reports are made concerning the mental status of individual pupils. To repeat: In the past when our inferences were confined to *averages*, we escaped these invalid conclusions. Now that we propose to speak of the *individual*, we must not overlook the statistical implications of the step that has been taken, but provide a technic that will make possible what we are attempting.

THE EFFECT OF KINAESTHETIC FACTORS IN THE DEVELOPMENT OF WORD RECOGNITION IN THE CASE OF NON-READERS[1]

GRACE M. FERNALD AND HELEN KELLER
University of California, Southern Branch

The cases reported in this paper are all those of children of normal mentality who have failed to learn to read after three or more years in the public schools. In all cases but one the vision was normal. The method described here was used only after the child had been given several weeks of individual instruction by recognized methods and had failed to make any improvement.

Many children who have been brought to us as non-readers with individual instruction and proper motivation, learned to read quite easily by ordinary methods when they were given individual instruction and proper motivation; others proved to be mentally deficient. In five years we have found only seven cases of actual non-readers, even though children have been brought to us from all parts of the state. In all seven cases the presumption of mental deficiency had been made as the explanation of the reading failure. In all but one case, however, the intelligence quotient was found to be at least 100 by the Stanford Revision.

METHOD

1. *Learning first words.*—The child was asked to tell some word he would like to learn. The word was written in large script on the blackboard or with crayola on cardboard. The child looked at the word, saying it over to himself and tracing it if he wished to do so. The tracing was done with the first two fingers of the right hand (or of the left hand if the child was left-handed) resting on the copy. It was never done in the air or with pencil. When the child was sure he knew the word, the copy was erased and he attempted to write the word, *saying the syllables to himself* as he wrote them. If he was unable to write the word correctly, the entire process was repeated until the word could be written without the copy. At no stage of the performance was he allowed to copy the word. After a few words had been learned in this way,

[1] This paper is to be followed shortly by two others, giving the results of experiments with first grade children and of experiments in spelling. The bibliography will be published with the final paper.

he was shown the word in print as well as in script. The next day he was shown the word in print only. If he failed to recognize it, it was written for him. If he still failed to recognize it, it was retaught as on the first presentation.

2. *Spontaneous sentences.*—After the first few days the child began to ask for sentences instead of words. A sentence was then written and he learned the words comprising it, finally writing the entire sentence as many times as he wished—always from memory, never from copy.

The sentences the child had requested were then printed on cardboard or were typewritten. These sentences and others, made of the same words, were read by the child. The same words were repeated in different sentences from day to day.

3. *Words in context or story selected by the child.*—As soon as the child was able to make out simple sentences, he was taken to the library and allowed to select a book. The first paragraphs read were worked over in the following manner. Before the reading, each word which had not already been learned was exposed through an adjustable slit in a piece of cardboard. If the child failed to read the word it was pronounced for him. He pronounced and then wrote the word (as before without looking at the copy). If he had difficulty in writing the word after seeing it in print, it was written for him and taught from the script as in the case of the first words.

4. *Apperception of phrases.*—After the words in the new paragraph had been taken up in this manner, brief exposures of the words were given until the child was sure of them. When recognition was immediate for every word, the slit was adjusted to phrases, and flash exposures of the various phrases were given. The exposures were never long enough to permit the phrases to be read word by word. As many successive exposures as were necessary for recognition were given. After the entire paragraph had been worked over in this way, the child was told to read the paragraph to himself and report what he had read.

5. *Silent reading for content.*—As soon as possible the child was encouraged to read to himself. There was no difficulty in any of our cases in getting him to do this after his progress had gone into the fourth phase as described above.

Description of Cases

Seven cases have been successfully treated to date. Four of these are described in the following pages. In one case, included under the above statement, the work is incomplete at the present time, but the results showed normal progress during the period of experimentation. Two additional cases, one from New York, and one from Colorado, now under investigation, have both reached the third stage. In the latter of these two cases the child is now able to learn new words rapidly from print and to write them correctly although he was unable to read or write his own name two months ago.

Case I. Boy (Lester)

Age (December, 1916) 10-2; mental age (Binet Simon, 1911 Scale), 9-0. Re-tested December, 1920; age, 14-2; mental age (Stanford Revision), 13-4; I. Q., 94. Vision, normal.

School history.—Lester had been in the public schools of Los Angeles for five years (including one year in the kindergarten). When brought to our notice he had been in the second grade of the Normal Training School for two months, having been entered in the Training School as a last resort after his failure in the city schools. On account of his size, he was first placed in the second grade; but, as he was unable to distinguish one-syllable sight words, he was put back into the first grade. Here a special effort was made to teach him to read in a small group where special instruction was possible. Since he made no progress whatever, he was reported to the psychology department as mentally deficient.

As the results of the mental test did not show sufficient retardation to account for his complete inability to read, we kept the boy under observation for some time. The case came to our notice at a time when we were particularly engaged with poor spellers. Accordingly, we put Lester into a group of our worst spellers, where individual work was being done.

Method.—The work in the spelling class referred to was being conducted in the following manner: The children watched while the word was written on the blackboard or on cardboard. They said the word over to themselves, tracing it with their fingers if they wished to do so. The tracing was done with the first two fingers of the right (or left) hand in contact with the copy. Tracing in the air or with pencil did not seem to produce the same results as tracing with the fingers. When each child was sure of the word, it was erased or otherwise removed from view, and the child wrote the word, saying the syllables to himself as he wrote them. He was allowed to write the word as many times as he wished, provided he did not copy it from a previous writing. Except for this special work in spelling, the method was the same as that described on page 355.

Results.—Lester developed a craze for writing words. He would work at it by the hour, tracing words over and over again. To our surprise,

after the first week, he seemed to be able to write from memory words which he had written previously several days in succession.

We had the same words which he had learned to write printed; and each time the boy wrote a word in script we showed him the corresponding printed form. Recognition was developed almost immediately for the words he had written. At first he learned only two or three words a week, then several a day. After about two months, he was able to look at a word in print, say it over to himself, and write it correctly. This was true even of words of three and four syllables. At this stage of his development, after he had once written a word, he would almost invariably recognize it on successive presentations. Yet, on the other hand, if told a word over and over again on successive days, he failed to recognize it unless he wrote it.

Since the boy was interested in history, and especially in the war, we chose a history of the United States as our first book. The method employed is described in the introductory statement on method.

Unfortunately we had no idea of the significance of the work and kept no records of the exact words written from day to day, as we did in our later cases. Much of the writing was done on the blackboard and immediately erased.

The boy seemed suddenly to go ahead by jumps. At the end of five months he was reading so readily that we had a demonstration before his various first- and second-grade teachers. They refused to believe that he was actually reading until they brought books selected by themselves and tried him out with both oral and silent reading.

After six months in the University Training School, the family moved and the boy went back to the regular city schools.

December 15, 1920, just four years after our first experience with Lester, we found him doing satisfactory work in the *seventh* grade of the city schools. His teacher was surprised when we asked about his reading, and said he was an inveterate reader.

TABLE I. RESULTS OF READING TESTS

Date	Test	Results					
Dec. 1916	Several	Zero on all tests; could not read or write words like "cat"					
Dec. 1920	Kansas	Lester's Results		Standard		Ability Shown	
	Silent, FormII, Test 2	Rate 106	Comprehension 24.7	Rate 106 (For Eighth Grade)	Comprehension 7th Grade, 23 8th Grade, 26.4	Rate 8th Grade	Comprehension Above 7th Grade

Note: Lester had never seen any of the upper grade Kansas Silent Reading Tests before the test was given December 1920.

Summary of case. At the age of ten years, after over four years in the elementary grades of the public schools, Lester was totally unable to read.

In six months he developed from a zero score in reading ability to ready word recognition. For the last three and a half years, without further special instruction, his development in both silent and oral reading has been perfectly normal. During this period he has advanced from the first to the seventh grade.

CASE II. BOY (HENRY)

Age, 9-8; Mental Age (Stanford Revision), 12-8; I. Q., 131. Vision, normal.

School history.—Henry entered the kindergarten of the University Training School when five years old, and has attended the same school continuously to date, except for part of one year lost on account of measles and whooping cough.

He seemed unable to learn to read or spell although he had been given individual instruction, including drill in phonics. He was referred to the psychology department as a possible mental defective. Although the results of the mental tests indicated unusual ability, he was found to be below second-grade rating in the Kansas Silent Reading Test and in the Ayres Spelling Scale. He also failed completely in all tests for phonics.

Method.—The method was the same as that previously described, except that Henry did not tend to trace the words. He would look at a word which had been written for him, say it over slowly, have the copy removed, and then write it saying each syllable as he wrote it. He soon developed considerable skill in writing words in this way; but he was very slow at first in making the association between the word he had written and the printed word. In the beginning it was necessary to present each word daily for five or six days before the association with the printed word was formed.

There was so little progress for the first few weeks that the psychology department was consulted repeatedly with reference to the seeming failure of the boy to develop word recognition. His mother came often to observe and expressed herself as convinced that it was simply another failure. She felt that she had tried everything. Besides being in a small class with a particularly successful teacher, he had received special help at home and special work in phonics with the reading department. Every effort had been made to encourage the boy; he had even been promoted to the fourth grade in order to try the effect of encouragement. As he could do nothing with fourth-grade reading and spelling he went back into the special study room. It was with difficulty that we obtained the mother's consent to have the boy tested mentally. She has since told us that she was afraid he would be found mentally deficient When informed of the results of the test, she was frankly skeptical. It should be stated, however, that she was very careful not to let the boy know of her discouragement and that she cooperated with us throughout the experiment.

After the fifth week Henry suddenly began to make such rapid progress that it was almost impossible to keep track of him. He could write any word after a single presentation and after he had once written it, he would recognize it in either script or print on subsequent presentation. As one of the student teachers expressed it, he was "on." Five weeks after our first work

with him, he wrote the word *Switzerland* after one presentation and recognized it in print twenty-four hours later.

Henry's first reading was in the second reader. It was necessary to write the new words and to have him write them before he could recognize the word on subsequent presentations. The development of a reading vocabulary was very rapid, however, after the first few weeks, since it was rarely necessary for him to write a word more than once, and since he was able to give one glance at the printed word, say it, and write it. He worked much by himself, asking anyone to pronounce a word, and then writing it off on a slip of paper.

From the time when Henry began working alone it was difficult to keep track of his progress. By May, three and a half months after we began the experiment, he was working out many new words without having them pronounced for him or without needing to write them. Though he had received no instruction in phonics since the experiment began, and had seemed at that time to have no knowledge of them, a test given May seventeenth showed the beginning of the development of the recognition of phonetic units.

Work was begun in February and discontinued in June at the close of the school year. The mother reported that the boy read everything available during the summer—library books, newspapers, advertisements, etc. On his return to school he went into the regular fifth grade. At the date of this writing (March, 1921), he is doing satisfactory work in all his subjects. Formal tests given in November showed his ability to be between that of the fourth and fifth grade according to both the Starch and Kansas tests, and above that of the sixth grade according to the Gray Oral Reading Test.

Formal reading, spelling, and phonics tests.—The following results were obtained from tests of spelling, reading, and phonetics. On January 30, 1920, Henry was given a selection of words from Column I of the Ayres Spelling Scale. His rating was ten percent, and the character of this performance may be judged from the fact that the second-grade standard for these words is fifty percent. The following are the word forms as he actually wrote them, the words in parentheses being the correct forms.

1. I (catch)	8. gon (gone)	15. bay (buy)
2. lake (black)	9. se (suit)	16. spo (stop)
3. wrme (warm)	10. trak (track)	17. wa (walk)
4. inles (unless)	11. w (watch)	18. grat (grant)
5. Cothing (clothing)	12. daas (dash)	19. soke (soak)
6. be (began)	13. fell (fell)	20. nu (news)
7. ababl (able)	14. fight (fight)	

On April 25, 1920, another selection of words was given from Column I of the Ayres scale. This time Henry's record was ninety-five percent which compared favorably with the standard for the fifth grade (ninety-four percent).

On February 6, 1920, and again on November 22, 1920, Henry was tested with the Kansas Silent Reading Test, Form I, Test I, with the following results. In the first test his rate score was 36 and his comprehension

score was 4. The lowest grade for which standards on the Kansas Silent Reading Test are available is the third grade. For this grade the standard in rate is 52 and in comprehension, 7.2. Thus, Henry's performance was considerably below that which one would expect of a typical third-grade child. On the second application of the test (November, 1920) Henry's rate score was 103 and his comprehension score was 18. This compares favorably with the fifth-grade standard of 89 for rate and 19 for comprehension.

On November 22, 1920, Henry took the Starch reading test for the third, fourth, and fifth grades. His scores in comparison with the standards for each of these grades follow.

Test	Henry's Score		Standard	
	Speed	Comprehension	Speed	Comprehension
3rd grade	2.3	33	2.1	20
4th grade	3.2	31	2.4	24
5th grade	2.6	21	2.8	33

Phonics test.—Although every effort has been made to teach Henry phonics in the second and third grades, the results of all phonics tests in January, 1920, were entirely negative. He had no idea of the sound significance of letter combinations.

The following results were obtained in tests given by the reading department. No instruction in phonics had been given since the experiment was started.

Test Given May 17th, 1920

1. *ash*—bash, hash, mash, crash, flash, nash
2. *in*—fin, sin, bin, lin, tin, men
3. *ed*—bed, led, he
4. *et*—let, set
5. *ind*—blind, sind, find, dind
6. *ong*—bing
7. *end*—fend, send, senll, mend
8. *ase*—base, mose
9. *am*—fan, sam, fan
10. *ite*—
11. *all*—fall, ball, mall, hall, tall, dall
12. *inch*—pinch
13. *ale*—pale, fale, sale, male, tale
14. *age*—page, cage, mage
15. *ad*—sad, mad, lad

Test Given December 24th, 1920

1. *hack*—rack, tack, mack
2. *all*—ball, tall, call, mall
3. *ate*—hate, mate, fate, late
4. *ent*—rent, ment, sent, tent
5. *ing*—ring, sting, bing
6. *ow*—wow, cow, row
7. *air*—hair, lair, rain
8. *ail*—mail, lail, fail, rail
9. *and*—hand, sand, cand
10. *ang*—rand, dang, sang
11. *ank*—sank, bank, flank, tank
12. *ark*—mark, hark, lark
13. *atch*—catch, match, atach
14. *eck*—reck, heck
15. *end*—attend, blend, afend
16. *ure*—sure, picture

Summary of spelling and reading tests.—(a) Spelling: Progress from below second grade to fifth-grade standing in three months. (b) Reading: Progress from much less than third-grade record to fifth-grade in six months. (c) Phonics: Although no instruction had been given in phonics between January and May, and although Henry seemed to have no knowledge of phonics at the earlier date, tests given at the later date showed the development of associations between many letter combinations and their sound values. Tests given in December of the same year showed his knowledge of phonics to be quite equal to that of the average child in the fifth grade.

Summary of case.—A year ago, Henry was not able to do second-grade work in reading and spelling. After four months' instruction, he was put into the fifth grade and has had no difficulty in keeping up to the fifth-grade standards for the five months of this year. He is now in the second half of the fifth grade, and is reported as reading books of every description out of school hours.

CASE III. BOY (FRED)

Age, 9-2; mental age (Stanford Revision), 9-3; I. Q., 100. Vision, right, normal; left, two-thirds normal.

School history.—Fred attended the public schools of Riverside, California, where he entered the kindergarten at the age of five. He spent three years in the first grade. He entered the University Training School, September, 1920, and was placed in the second grade but made no progress during the first month. He was unable to read words of one syllable, and consequently was too poor to be tested by any formal test. He was sent to the Psychology Department for a mental test with the definite suggestion that he be transferred to one of the public school rooms for the mentally deficient. The child had so much the appearance of a mental defective that we were surprised at the results of the mental tests.

A month after his entrance into the University Training School Fred was given tests in phonics and in spelling. In phonics he wrote such unrelated forms as the following: for *on, am;* for *at, oot;* for *in, tn;* for *ish, ih.* In spelling he missed all but two of eight words, writing the following forms for the words given in parentheses:

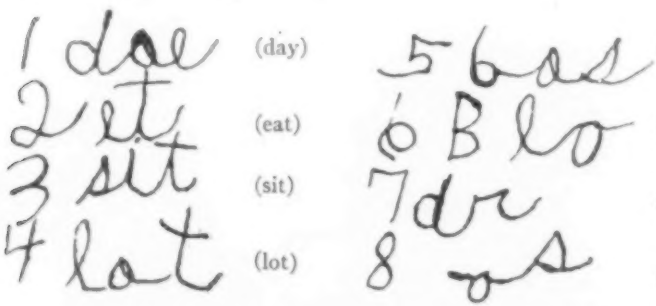

FIGURE 1. WORDS WRITTEN BY FRED SEPTEMBER 29, 1921

First phase—learning first words. Fred spent the morning learning to write the four words *will, you, am, boy,* all of which he had asked for. The words were taught one at a time; each being written by the teacher on the board. Fred said each word over and traced it with his fingers until he was sure he could write it. It was then erased, and he attempted to write it. If the first attempt was not successful, the entire process was repeated until the word was written correctly.

Fred succeeded in writing each of the words correctly in the course of the morning. In the afternoon he attempted to rewrite the same words with the following results: *W* (for will); *Yo* (for you); *ow* (for am); *b* (for boy). The words were retaught and then written correctly. He then asked for the word "I," and wrote:

I am o boy

He wrote this sentence with his left hand, then with his right. There is little difference in the performance with the two hands, but Fred says he likes the left hand better. He usually writes on paper with his left hand and on the board with his right.

TABLE II. RECORD OF FIRST WORDS LEARNED BY FRED

Date (October)	Word	Number of Times Presented	Date (October)	Word	Number of Times Presented
6	will	2	7		1
7		0	7	yes	2
6	you	5	8		1
7		2	25		0
8		1	8	last	3
25		0	10		2
6	am	3	11	little	1
7		1	12		0
8		0	11	Pine	2
6	boy	1	12		0
7		0	11	tree	1
7	door	2	13		0
7	box	3	11	grew	5
7	open	1	13		0
7	mouse	4	11	woods	5
7	and	2	13		0
8		2	12	October	2
8		2	13		0
25		0	12	of	1
7	we	1	13		0
7	day	3	12	words	2
7		1	21		0
7		0			

October seventh.—He wrote correctly from memory the four words which he had learned the previous day, then asked for *door* and *box*. After working for some time on these words, he asked for *open*, adding "then I'll learn *the* next and I can write *open the door*." When the word *open* was written by the teacher, Fred said, "Erase it quick so I can write it." He next asked for the sentence *Open the box till the mouse jumps out.* The word *mouse* had to be presented four times before Fred could write it correctly. By "presented" is meant that the word was written by the teacher and traced by the pupil until the pupil felt sure he could write it.

Table II gives the record of the first words learned by Fred and the number of times it was necessary for each to be presented before he was able to write it correctly. Table II also shows the results of attempts to write the same words on later dates. The words were all asked for by Fred.

For the next three weeks Fred went through an orgy of writing. He wrote blackboards full of words, then began to write sentences, and finally wrote the letter to his father, as shown in Figure 2.

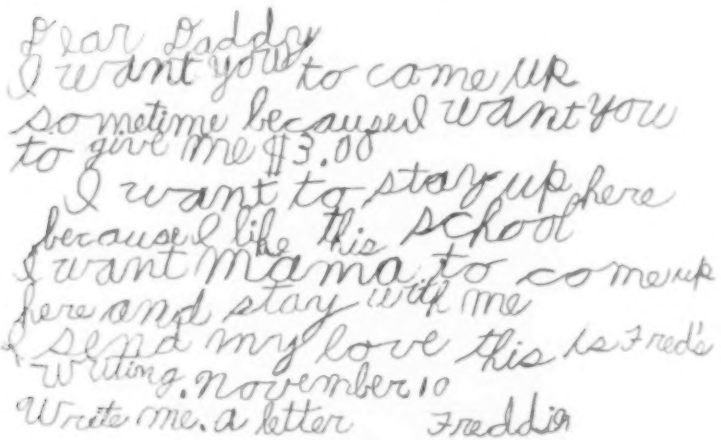
Dear Daddy
I want you to come up
sometime because I want you
to give me $3.00
I want to stay up here
because I like this school
I want Mama to come up
here and stay with me
I send my love this is Fred's
Writing. November 10
Write me a letter Freddie

FIGURE 2. LETTER WRITTEN BY FRED WITHOUT ASSISTANCE NOVEMBER 10, 1921 (THE WORDS IN THE LETTER HAD BEEN PREVIOUSLY ASKED FOR AND LEARNED)

This letter is only one specimen of the spontaneous compositions with which the boy kept himself occupied by the hour. It was written six weeks after the attempt to write the words given in Figure 1. He was constantly asking for new words which he learned and wrote from memory. For over a month after writing his first sentence, Fred seemed to care little about subject matter, so absorbed was he in mastering the mechanics of writing. He worked so constantly and so hard that it was necessary to force him to leave the room at recess and at the close of school. He looked up one day after working for two hours at new words and exclaimed, "You know I

scarcely ever used to get promoted and now just look at all I am learning." As soon as we attempted to teach him words without having him write them, his interest was lost. He would try to learn them by saying them over and looking at them, but would soon become discouraged and would fail to recognize the words after repeated presentation. This attitude in the early stage of the experiment is particularly interesting, because at the present time (March, 1921) Fred no longer wishes to write unless he has something special to say, but is reading everything he can get hold of. He is as eager over making out new words without writing them as he ever was over the writing process.

Second phase—The association of written with printed words.—Words for which Fred asked were written for him on the typewriter or were shown him in print after he had written them. Simple reading exercises were given, using the words he had written.

Table III shows the results with the first words studied for the purpose of reading. The printed word was first shown the boy and then written for him. He studied it on first presentation as he did the words in Table II, but he was not asked to write a word a second time unless he failed to recognize it on later presentations. Column 3 shows the number of presentations necessary before the word was written correctly.

TABLE III. RESULTS OF FIRST WORDS PRESENTED FOR READING PURPOSES

Date of first presentation (Oct. 1920)	Word	Number of presentations	Date of first presentation (Oct. 1920)	Word	Number of presentations
14	pussy	2		build	3
	where	3		tired	3
15	fur	4	20	month	3
	would	4		sleep	1
	rings	2		dinner	1
	thanked	3	21	chair	1
	fields	1		bowl	2
	over	1		middle	2
18	sunshine	2		bears	1
19	great	2	25	heard	1
	name	1		across	1

Third Phase—Writing the word from memory after looking at the printed copy and having the word pronounced.—Within six weeks after the experiment was started, it was never necessary for Fred to have the word written for him. He was able to look at printed words of several syllables, say them over to

himself, and write them correctly from memory. On November seventeenth he wrote *disappointed*, *department*, *university*, and *training* after seeing the words once in print. A week later he recognized all of these words without any hesitation, and without having seen them in the interval.

Development of projects in connection with reading and writing. November seventeenth.—Fred suddenly decided that he wanted to draw, and was allowed to do so as long as he wished. He drew a picture of a cannon and then of a house and garage with heating system. His picture of the garage is shown in Figure 3. It was then suggested that he label his pictures, and this idea pleased him greatly. He learned to write the following seven words and used them as labels: *cannon, shells, pipe, steps, window, garage, furnace.*

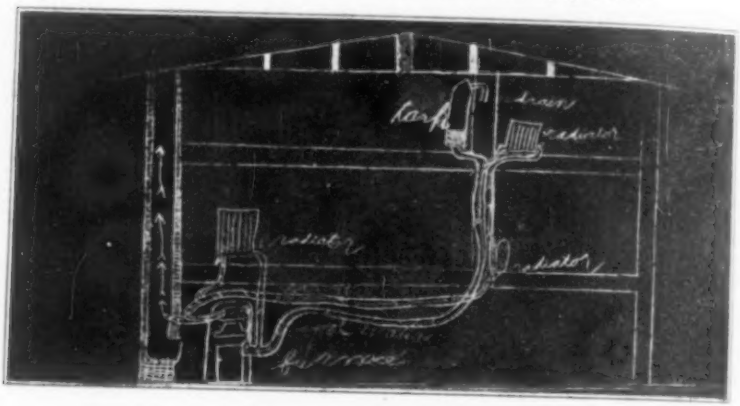

FIGURE 3. PHOTOGRAPH OF ONE OF FRED'S DIAGRAMS

November eighteenth.—He wrote correctly without presentation the following words which he had previously learned: *you, ride, bicycle, furnace, hide, pipes, rope, because, garage, cannon, want, Riverside, December, October, will, live.* He then wrote above the words on the blackboard, "These words are saved."

November nineteenth.—He was taken to the library and allowed to select two books: Lynde's *Physics in the Household*, and a book on plumbing. From November nineteenth till December sixth, Fred drew and labelled diagrams, finding the words in physics and plumbing books. He was still unable to recognize new words, after he had been told what they were, unless he wrote them. One writing even of very difficult words was usually all that was necessary for word recognition.

November twenty-second.—Fred read the following paragraph without being told any words except those enclosed in parentheses:

> The cooler water in the radiator (being heavier) sinks from the radiator into the furnace (boiler) and (forces) the hot water from the (boiler) into the radiator. This hot water gives up its heat to the air in the room and (thus) cools (contracts) and becomes heavier. It then sinks back into the boiler and forces more hot water into the radiator.

This paragraph was read just one and a half months from the time when the boy could not read the simple sentence "*I am a boy.*" He stumbled over a simple word like *thus* which he had never written, but had no difficulty in reading any of the much harder words which he had written. He was told *thus* each time he failed to read it, but was quite unable to recognize it on later presentations until he had written it, after which he recognized it whenever it occurred.

December sixth.—Fred suddenly stopped drawing and labelling diagrams and began writing on the blackboard sentences which he would sometimes illustrate with pictures. These sentences and others, containing the same words, were then typewritten and given him to read. For example, the following sentences with a picture of a racing machine at the top of the sheet were written after a visit to the automobile machine shop.

A RACER

This is a fine automobile.
It has four wheels on it. They have tires and mud guards.
It has a steering gear, and a crank on the front.
It has a windshield. It has a radiator.
It is a Mack and it is a racer.
It runs awfully fast.
Gasoline runs the engine.
The radiator has water in it to cool the engine.
The engine has eight cylinders.
This machine has United States tires.
The engine is a fine one.
The fan helps keep the engine cool.
This machine can turn corners going very fast and it won't wreck.

December 13, 1920.

The following 206 words, written from October sixth to December twenty-second inclusive, are arranged in the order of the frequency with which Fred used them. The numeral after each word indicates the number of times it was used.

On December 22 all the words in this list were shown to him and he recognized all of them except the four marked with an asterik.

I....25	water....11	because....8	me....5
these....22		cannon....6	on....5
a....18	want....10	has....6	Riverside....5
		pipes....6	words....5
and....14	it....9	radiator....6	
this....14			come....4
the....14	cool....8	can....5	draw....4
to....14	furnace....8	day....5	fine....4
	ride....8	engine....5	fur....4
like....13		garage....5	little....4
	hot....7	hide....5	October....4
is....11	will....7	live....5	*page....4

saved....4
tired....4
where....4
would....4
yes....4

bicycle....3
box....3
boy....3
build....3
car....3
chimney....3
crank....3
December....3
door....3
fast....3
*month....3
mouse....3
Mrs....3
rope....3
stay....3
up....3
with....3
write....3

are....2
arithmetic....2
autumn....2
cold....2
have....2
heating....2
in....2
last....2
love....2
machines....2
mama....2
now....2
of....2
open....2
pipe....2
pussy....2
read....2
rings....2
sun....2
saw....2
school....2
send....2
sunshine....2
system....2
tank....2
*thanked....2
tires....2
umbrellas....2
university....2
we....2
wires....2
window....2
wood....2
won't....2
wreck....2

across....1
air....1
an....1
as....1
auto....1
automobile....1
awfully....1
ball....1
bears....1
beat....1
belong....1
birthday....1
blew....1
board....1
boat....1
bought....1
boiler....1
bowl....1
bridge....1
buttoned....1
buy....1
cackle....1
carry....1
chair....1
circus....1
closer....1
coming....1
corners....1
cylinders....1
daddy....1
dear....1
department....1
didn't....1
dinner....1
disappoint....1
drew....1
easy....1
eight....1
expect....1
fan....1
faster....1
fields....1
fireplace....1
four....1
front....1
garden....1
gasoline....1
gear....1
give....1
going....1
grew....1
heard....1
help....1
here....1
holiday....1
house....1
hurt....1
*journey....1
letter....1
lot....1
lazy....1
middle....1
motor....1
mud guards....1
my....1
name....1
never....1
November....1
one....1
over....1
pin....1
pine....1
play....1
plough....1
poles....1
racer....1
railroad....1
riding....1
rose....1
run....1
runner....1
shells....1
shoe....1
sign....1
sit....1
sky....1
sleep....1
sometime....1
spark plug....1
stand....1
steering....1
steps....1
tail....1
than....1
that....1
track....1
training....1
tricks....1
trolley....1
turns....1
unable....1
under....1
United States....1
very....1
walk....1
warmer....1
wheels....1
wire....1
writing....1

All but four of the words in the above list were asked for by Fred and learned because he wished to read or to write about some particular thing. The four exceptions were "department," "disappointed," "university," and "training" which were given to demonstrate the readiness with which he could learn new words.

The numbers after the words indicate the number of separate occasions upon which the word was written in context. The boy often wrote a word over and over again voluntarily, erasing it each time it was written, but as no record was kept of this the actual number of times each word was written is much greater than indicated.

December twenty-second.—Fred began to ask for new words and to remember them without writing them. *December twenty-fourth—January fourth.*—Christmas vacation.

Fourth phase—Ability to pronounce new words if they resemble words already written. January twenty-fifth.—Fred sounded out the word *mother*, then said it over several times. He was quite excited over being able to say the word without being told, and began to attempt the same thing with other words.

February fourth.—He worked out, without writing or any help, the following words: *surprised, roar, fright, noise, thirsty, dirty, middle.* He had to be told *dreadful*, pronounced *towards, to-wards*, and *vines, vī-nēs*. He read the fable of the *Fox and the Lion* to himself and told the story, giving every detail. He had never heard or read the story before.

March third.—Words were given to Fred from the pages of the third reader. The longer words on various pages were given out of context to see whether he could read them without having them pronounced. All the words were new to Fred so far as we know. He read all the words in the following list, mispronouncing only those which are starred. In each case the mispronunciation is indicated.

enough	money	roasting	*ōnly
crowing	with	ashes	wounded
screamed	tripped	quite	howled
tumbled	*wârmth	loaded	stolen
midnight	frozen	lazy	empty
*chămber	traveler	healthy	shoulder
*quarreling	chestnut	branches	*cruelly
quarrel	graceful	whining	brooks
mistress	*grīnned	*răttling	California
maid	monkey	apple	state
*dŭbled	second	grave	series

At the date of writing this paper (March 1921), Fred's progress is so rapid that it is difficult to keep records. He takes library books home, reads to himself, and has to be told only such words as might trouble any child of his age. He has been in the regular third grade for two weeks and is having no difficulty with the work. If his progress continues at the present rate, he should be able to make up the one year necessary to put him in the grade appropriate to his chronological age.

Summary of case.—In October 1920, Fred could not read or write even monosyllabic words. He failed completely in all tests for reading, spelling, and phonics. His school report showed steady attendance with failure of promotion in the city schools of Riverside, California. It was taken for granted that he was mentally deficient until mental tests proved him normal.

At the beginning of the experiment his progress was very slow. He seemed wholly dependent on tracing the words first learned and continued to trace for over two months. His development went through well-marked stages, with the transition from one stage to the next apparently quite sudden.

In March 1921, five months after the experiment was started, he was reading and writing (spelling) well enough to go into the regular third grade. He was taking library books home to read to himself and could read any ordinary story and give its content.[2]

CASE VI. (JOE)

Age (June 1919), 12-5; mental age (Stanford Revision), 15; I. Q., 120. Vision, normal.

Although incomplete, this case is reported because it differs in certain respects from our other cases. *In the first place*, in spite of the fact that the boy showed almost no ability in reading or spelling, he did not have to go through the first two stages described on page 355. From the very start he could write a word from memory, after seeing the printed word, having it pronounced for him, and pronouncing it himself. Thus, he was able to begin with what had proved the third stage in all our other cases. *In the second place*, the experiment was discontinued before the fourth stage was reached. Joe had learned many new words so that he could read fairly well on certain topics; the work with rapid apperception of phrases had been begun; but the stage had not been reached in which new words were apperceived on the basis of their resemblance to words already learned. The progress of the case during the year and a half of regular school work since the experiment was discontinued is of interest in comparison with the progress made by cases in which the experiment was completed.

School history.—Joe has attended the Los Angeles city schools since he was six years old. From the start he has seemed to be unable to learn to read but has been passed from grade to grade on account of his ability in other subjects, particularly in arithmetic. His grades in the fifth year ranged from *B* in arithmetic through *C*− in geography and history to *D* in English. When questioned as to how he kept up the reading end of some of his school subjects, he explained that he had always managed to work with a boy who could read easily. He helped the other boy with problems in return for help in reading.

Joe was referred to us in June 1919 as a sixth-grade failure. Investigation of the case showed that the entire difficulty was due to an almost complete inability to read. There had been considerable discussion of the case before it came to the psychology department, since two other children in the same family had had the same difficulty with reading.

The case was the more surprising because of the boy's high intelligence quotient and because of his history outside of school. He excelled in games like baseball and was very popular with other boys. The proprietor of a

[2] At the date of publication, December 1921, Fred is doing satisfactory work in the upper fourth grade.

drug store where he worked out of school hours reported that he could keep track of things better than most men.

Since the case came to our attention at the end of the school year, we transferred the boy to a summer school where the experiment could be started. He attended this school intermittently for a month. Since then he has had a few hours of special work, extending over a period of about two weeks. He entered the seventh grade in the fall and is now (a year and a half later, March 1921) in the eighth grade in one of the regular Los Angeles city schools, where he is reported as doing satisfactory work.

Method.—The method was the same as that already described except that, like Henry, Joe never tended to trace words, and was able from the start to learn them directly from the printed copy. Consequently in the case of each word it was only necessary to pronounce it for him, to have him pronounce it to himself as he looked at it, and then to have him write it from memory, saying each syllable as he wrote it.

Results.—Although the results of the first reading and spelling tests were almost entirely negative, Joe possessed from the beginning a remarkable ability to learn new words by the method just described. In spite of the fact that he was unable to recognize short, ordinary words unless they were pronounced for him, he was able to look at a difficult word, pronounce it after someone, and write it correctly from memory, provided he said each syllable to himself as he wrote it. It seemed to make no difference whether he knew the meaning of the word or not. He could write a word like *psychophysical* after seeing it once and repeating it after some one, and could often rewrite the word, as well as recognize it, after an interval of several days. His ability to write words in this way was so unusual as to attract general attention. It became one form of school entertainment to try to find a word which Joe could not write. To this day he knows certain outlandish words learned in this way. It was, of course, not a part of the plan of instruction to teach any but ordinary words. During the five weeks of summer school he went over the Ayres thousand-word list and developed a considerable reading vocabulary.

A study of Joe as he works shows that he is more dependent than any of our other children on saying the word as he writes it. He makes marked lip movements as he says the syllables to himself and fails completely if he is made to suppress the lip movements.

He usually remembers for an almost indefinite period any word he has ever written, but fails on a subsequent presentation to recognize a word which he has been told repeatedly but which he has not written. At the end of the summer school, when the experiment was discontinued, he had reached the stage where he was beginning to make out a few new words on the basis of their resemblance to words he already knew. For the most part, however, he was still not able to recall words unless he wrote them. Work in rapid apperception of phrases had just been begun.

In the fall of 1919, Joe went into the regular seventh grade and did fairly well in all his studies, except English in which he failed. He seemed to have developed sufficient reading ability to keep up with his work in other subjects.

The experiment was discontinued because his teacher, since he was able to do the work in most of his subjects, considered it unnecessary to take time for the special reading work.

In March 1921, we found Joe in the eighth grade of the regular Los Angeles city schools. He was reported as doing satisfactory work, although slow in subjects requiring reading. Actual tests for reading and spelling give sixth-grade average for spelling and very irregular results for reading. He grades anywhere from the third to the sixth grade by the Kansas Silent Reading Test, below fourth grade by the Starch Silent Reading Test, and makes only half the standard score for the eighth grade by the Gray Oral Reading Test.

The discrepancy in the results of the silent reading tests is easily explained as soon as one observes Joe while he reads. He makes marked lip movements, and stops when he comes to some word he cannot recall, or else mispronounces some word to himself and so loses the sense of the whole. In the latter case he begins again at the beginning and either reads until he gets the meaning or gives up. For example, in the Starch test for the fourth grade, he missed the word *waked*, read till he came to *dawn*, went back to the beginning twice, and finally gave it up after getting *waked* and missing *dawn*. His score for Test 5 was much better than that for Test 4 because he did not happen to stumble over any particular word.

He is still able to learn a new word almost as rapidly as he can look at it and say it to himself. To determine whether the writing of the word is still essential for word recognition, the following experiment was tried:

He wrote the following words from paragraphs 8, 9, and 10 of the Gray Oral Reading Test: *dignifying, station, position, securing, approximately, scrupulously, inclined, contemptuous, silence, complexion.* Although he had failed to read any of these words, he wrote each of them correctly after a single presentation consisting only in the exposure of the word while it was pronounced and repeated once. The two words, *ingratiatingly* and *Josephus*, were the only ones on which he failed.

The words *proportioned, exegencies, profusion,* and *exhausted* were told to him four times and repeated by him each time, but he was not allowed to write them.

He was then asked to read the paragraphs again. He read correctly each of the words he had written but failed on all of the four words which had been pronounced but not written.

The results of educational tests given both at the beginning of the experiment and in March, 1921, indicate the amount of improvement which Joe made during the interval. These tests covered the subjects of spelling and reading. In spelling, words from Column Q of the Ayres scale were given on June 3, 1919. Joe's rating was 10 percent, which may be compared with the standard rating of 58 percent for the fourth grade on words of this column. On June 9, 1919, words from Column Q of the Ayres scale were again dictated. This time, however, the test was given after the words had been learned. Joe's rating was 90 percent. On March 7, 1921, he was tested with words from Columns Q and S of the Ayres scale. On Column Q

his rating was 85 percent (standard, 84 for the sixth grade); while on Column S his rating was 80 percent (standard, 73 for the sixth grade). In this connection it may be noted that his school report for March 7, 1921, indicated that his spelling was "good."

The reader may judge the amount and character of Joe's difficulties with spelling by consulting the following list of words from Column Q of the Ayres scale as Joe actually wrote these words on June 3, 1919. In each case the word in parenthesis is the word attempted.

somtines	(sometimes)	dclaer	(de lare)	ingae	(engage)
factoer	(factory)	finel	(final)	cention	(connection)
troruble	(terrible)	frem	(firm)	serprise	(surprise)
regen	(region)	prived	(period)	convick	(convict)
addtion	(addition)	privet	(private)	imploy	(employ)
proped	(property)	debate	(debate)	sclet	(select)
proin	(crowd)				

Table III shows the results at the beginning and end of the experiment according to the Gray Oral Reading Test. In June, 1919, his score was 7.5; in March, 1921, his score was 28.75.

TABLE III. JOE'S RECORD ON THE GRAY ORAL READING TEST

Paragraph	June 1919			March 1912		
	Time (seconds)	Errors	Score × Value	Time (seconds)	Errors	Score × Value
1	40	5	0	20	0	20
2	40	3	10	20	0	20
3	35	0	20	25	0	20
4	45	6	0	29	0	20
5	80	8	0	25	3	10
6	85	10	0	30	3	10
7	190	8	0	29	2	15
8	.....	.....	.....	90	7	0
9	.....	.....	.....	90	10	0

In June, 1919, Joe was quite unable to negotiate the Kansas Silent Reading Test for the sixth, seventh, and eighth grades. He was not given Test I (i.e., the test for the third, fourth, and fifth grades). In March, 1921, he was given both Test I and Test II. On the former he secured a rate score of 61 and a comprehension score of 9. On the latter he secured a rate score of 54 and a comprehension score of 13.

In the Starch Silent Reading Test he obtained the following results in March, 1921: on the fourth grade test, 0.45 words per second, comprehension 11; on the fifth grade test, 1.2 words per second, comprehension 19. An explanation of the inversion shown here is given elsewhere.

Summary of Spelling and Reading Tests.

June 1919. *Kansas Silent Reading Test.* Failure on sixth-, seventh- and eighth-grade form.

Gray Oral Reading Test. Score 7.5 in comparison with 49 as sixth-grade standard. Below first grade if computed on basis of first-grade test.

Ayres Spelling Scale, Column Q. Score 10 percent, standard for fourth grade 50 percent.

March 1921. *Kansas Silent Reading Test.* Grade from third to sixth, varying with test forms.

Starch Silent Reading Test. Below fourth, irregular.

Ayres Spelling Scale. Sixth grade.

Summary of case.—In June 1919, Joe was reported as a sixth-grade failure. Investigation showed that the entire difficulty could be traced to failure in reading and writing (spelling). It was found that he could not read ordinary material such as a simple arithmetic problem, though he could easily solve the problem if it were read to him. He failed almost completely in formal reading and in spelling tests. In spite of his inability to read or spell, the boy was able to learn new words with remarkable speed and accuracy by the methods already described.

After six weeks of rather irregular instruction, he had developed a considerable reading vocabulary along certain lines and had been brought up to the sixth-grade standard in spelling. He had improved sufficiently to be able to read his arithmetic problems and to do the work of the seventh grade, though he still read very slowly and stumbled over new words.

During the year and a half since the experiment was discontinued he has made little progress in ability to read to himself. He seems, however, to have lost none of the concrete detail he acquired, and he still manages to read fairly well in those subjects for which he learned a specific reading vocabulary. His spelling continues to be satisfactory because he learns to write new words easily, and is allowed to use the method developed during the experiment. Unlike our other cases in which the experiment was completed, Joe has never shown any tendency to read to himself and so has not acquired speed and ease in reading.

General Conclusions

In all but one[3] of the cases studied, progress seems to have taken place in four distinct phases, as follows:

Learning to write words.—In all cases the children were at first lacking in ability to write words as well as in ability to read. The development of ability to write words is very slow at first. It is necessary for the child either to trace or articulate the word many times while looking at the written copy, and finally to articulate it as he writes it from memory. The need for

[3] Case IV did not go through stages 1 and 2 but began at stage 3.

tracing gradually disappears, but he continues indefinitely to articulate in learning to write a new word.

Associating the written with the printed word.—The child sees the word in print, has it written for him, and then writes it himself, often tracing difficult words before writing them. He soon reaches a point where he can generally recognize a word in print after he has written it. He must still have the word written for him before he is able to write it himself.

Ability to write a new word from memory after looking at the printed copy and repeating the word to himself.—The word must, of course, be pronounced for him before he is able to say it to himself. He is still unable to recognize short, easy words on subsequent presentation if they are taught him in the usual way and if he does not write them. At this stage he will often write from 25 to 50 words a day. He rarely fails to recognize a word after he has once written it.

Ability to pronounce new words if they resemble words he has already learned.—The end of this stage is normal ability to read. The progress at the end is so rapid that it is almost impossible to keep track of the child's development. He seems suddenly to read and is able to enter regular classes in all work involving reading. In the first three cases the children not only developed normal ability to read, but became incessant readers.

Effect of intelligence on method and learning rate.—The method of learning was practically the same with cases of varying degrees of intelligence, except that there was no tendency to trace in the case of the two children with the highest intelligence quotients. In all cases the articulation and the writing of the word seemed essential for developing word recognition. The progress was much more rapid in the cases of better mentality than in those of lower mentality.

Persistance of kinaesthetic factors.—Children who have to trace words in the early learning stages continue to make slight hand and arm movements in attempting to recall difficult words or to learn new words. All the children make marked movements of articulation during the process of learning a new word, even after reading has been well developed.

Acquiring skill in penmanship and phonics.—Although there has been no drill in penmanship or in phonics, the children who

trace words in their early learning stages write a clear, free hand; and all acquire incidentally a good working knowledge of phonics.

In all of our cases any digression which directed the child's attention away from the word itself seemed to confuse him rather than to hasten the learning process. The introduction of phonics, formal penmanship drill, oral spelling, or even spoken directions, during the writing of the word seemed to hamper the learning process.

Individual differences.—The cases studied differed somewhat in the exact kinaesthetic content necessary for the development of word recognition, the difference being in the amount of hand and arm kinaesthetic experience necessary before the word was written. Case IV did not go through the first and second stages as described above. Although, in this case, the word was not traced before it was written and did not have to be presented in script, it was necessary for the child to write the word before he could recognize it.

General significance of results.—It may be well to note here that, although only extreme cases are reported in this paper, the results of certain experiments now under way seem to indicate that these general principles hold true of many cases in which the child simply has *difficulty* in learning to read. It seems that, at least in many of these cases, the progress will become normal if the proper kinaesthetic content is supplied.

Theoretical

Perhaps we can go no further in theory than to say that, in the specific cases studied, lip and hand kinaesthetic elements seem to be the essential link between the visual cue and the various associations which give it word meaning. In other words, it seems to be necessary for the child to develop a certain kinaesthetic background before he can apperceive the visual sensations for which the printed words form the stimulus. Even the associations between the spoken and the printed word seem not to be fixed without the kinaesthetic links.

The motor tendency is still obvious after the children become fluent readers. They seem far to outclass other children in the same grades in their ability to look at new words, say them to themselves, and write them. All of these children still make pronounced lip movements of saying the words (not the letters)

when learning a new word, even after they have reached a point where they never trace a word or speak it aloud. The children who traced in the beginning tend to make arm and hand movements in learning a new word or attempting to recall a difficult one. They are hopelessly confused as soon as they attempt to spell orally or to write a new word without saying it to themselves.

It would seem that the methods of teaching reading have always neglected the kinaesthetic factors, except those which in no way express the word as written or printed. It has been taken for granted that, in the case of all children, the visual cue is adequate to arouse those associations which make this cue stand for word meaning.

AN EXPERIMENT CARRIED ON WITH THE PUPILS OF THE RUSSELL PREVOCATIONAL ROOM

JAMES H. VOORHEES
Principal, Lynch School, Detroit, Michigan

The Prevocational Room of the Russell School consists of about fifty pedagogically retarded boys who have been weeded out of the regular grades through the Psychological Clinic and placed by themselves as a select group. The fact that the elimination of this type of child from the regular grades affords a great relief to the school system is axiomatic, but the purpose of this experiment was to find whether the proposition as it now exists can be defended on any other terms, that is, whether these boys show any growth in academic knowledge? Do they profit by the academic and manual instruction received here to the extent that they are better able to cope with life as they face it? It is not to be expected that any investigation carried on with a single group will speak conclusively for the various classes of this type in the City of Detroit. It is hoped, however, that in view of the growing importance of the problem the present work will be extended by others who will verify, repudiate, or modify the conclusions herein reached.

The present investigation, it will be seen, divides itself naturally into two distinct phases: first, the measurement of growth or gain in knowledge of the academic subjects pursued in the room; second, a follow-up study for the purpose of finding the exact situation of those who had gone from the room to become members of their respective communities. To determine the amount of growth in the subjects pursued by the boys, the following Detroit Standard Tests were given: spelling, arithmetic, geography, writing, and the Trabue Language Scale. The initial tests were given on separate days during the first week in January, 1919, and the final tests were given during March. Figure 1 shows the results of these scores. It is to be noted that these tests were identical in every case and that the period between the initial and the final tests was about two and a half months. The arithmetic test was a test of the four fundamental operations.

A gain of 16.5 percent was made in spelling, and a gain of 6 percent in arithmetic. The class apparently had not profited from the instruction received in the other branches during this

ten-week interval. The loss sustained in city location may have been due to the fact that but little drill had been given on cities during the intervening period. The small variation in the scores made in handwriting (not shown in the figure) language and state location tends to indicate that each test was a fair representation of the individual abilities and that their limit had been reached.

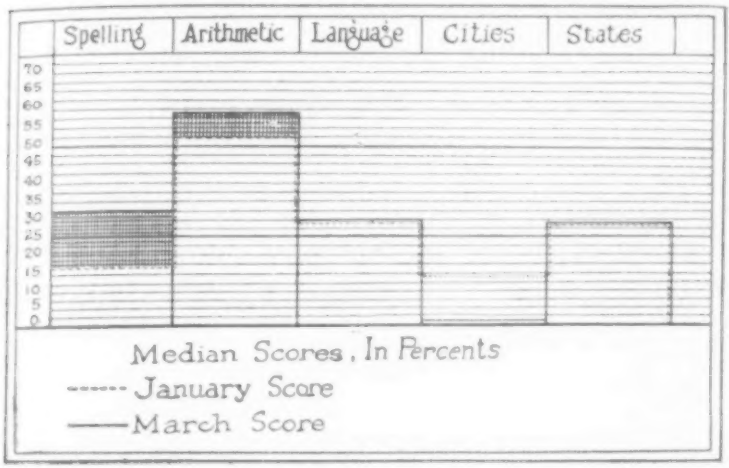

FIGURE 1. GROWTH IN ABILITY MEASURED BY CERTAIN TESTS

A second experiment in academic measurement was carried on during 1919–1920 in the same room. This experiment was an attempt to determine the relative growth in academic knowledge of the Prevocational Class, of a class of the same mental age, and of a class of the same chronological age, all within the same building. A class of B-third-graders was selected as representatives of the same mental age, and pupils of the grade who were at age, neither retarded nor accelerated, were chosen because the average mental age of the prevocational group was nine years. A class of B-eighth-graders who matched these boys in chronological age was also selected. These three groups were given identical tests at the same time in December, 1919, and in May, 1920, the interval being about six months. It was apparent in tabulating the results of these tests that the third graders were the best to use for comparison, the eighth graders having registered too high a score in the initial tests even though the time allowance was adjusted. The

time allowance for the third graders and for the prevocational class was the same in all tests.

This experiment seemed the more interesting because of the longer intervening period between the giving of the tests and because of the fact that the prevocational boys were given a different teacher during five of the six months that the experiment ran. The tests consisted of the following Detroit Standard Tests: Arithmetic Test No. 16, an addition test; Arithmetic Test No. 6, a reasoning test; Test 1B, a sentence organization test; a spelling test, made from the Ayres scale; a writing test, measuring quality and rate; and a silent reading test, measuring rate of reading and index of comprehension. All of these tests were identical except the spelling test, which was devised from the Ayres scale by selecting different words of the same difficulty for the initial and the final tests. The results of the handwriting are shown first. It is well to note here that a program was arranged by the teacher whereby these boys were given a twenty-minute lesson in penmanship two and three times a week for a four months period by the assigned penmanship instructor of the school. The results of this experiment are compared also with the results of the year before when practically no attention was given to penmanship as a subject. Figure 2 shows the scores made in these tests.

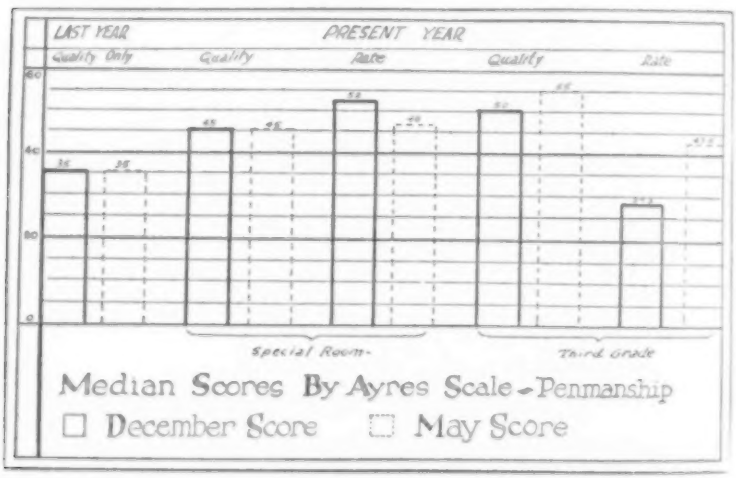

FIGURE 2. COMPARISON OF GROWTH IN PENMANSHIP OF A PREVOCATIONAL AND A THIRD-GRADE GROUP

The same quality of handwriting was shown in each test of the first year, and in the scores of 1919–1920 the same quality was attained in each test while the rate had slowed down six words. Evidently the class had not profited by the four-months period of efficient penmanship instruction. The experiment was very much worth while, however, in view of the added light thrown on the investigation. The third grade over the same period of time showed a gain in quality from fifty to fifty-five, and in rate, from twenty-seven and three-tenths to forty-three and five-tenths words per minute. The quality was ten points higher than that of the prevocational boys, and a good gain was made in rate of writing.

Figures 3 and 4 show the results of the scores made in arithmetic, sentence organization, and spelling by each class.

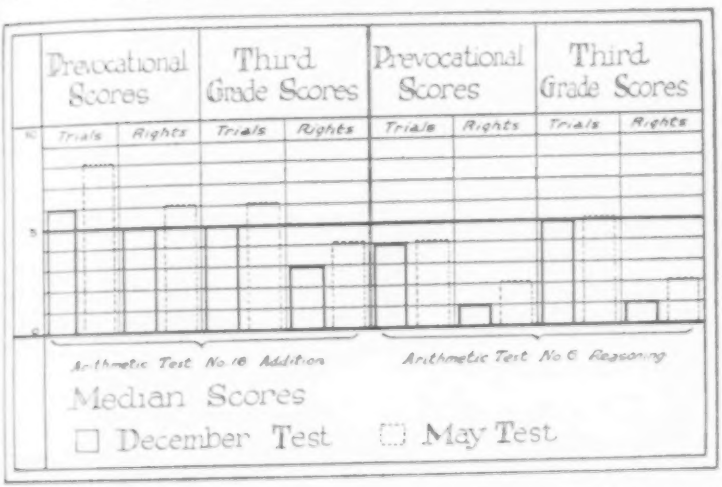

FIGURE 3. COMPARISON OF GROWTH OF A PREVOCATIONAL AND A THIRD-GRADE GROUP MEASURED BY TWO ARITHMETIC TESTS

The outstanding feature of these graphs is the great similarity between the gains of the two groups, and the low scores of the prevocational group. For example, in the addition test the prevocational boys had profited by this six months interval to the extent that they were able to attack two more problems, the third graders one more; the boys were able to solve correctly one more problem in addition, the third graders one more; in the reasoning test both

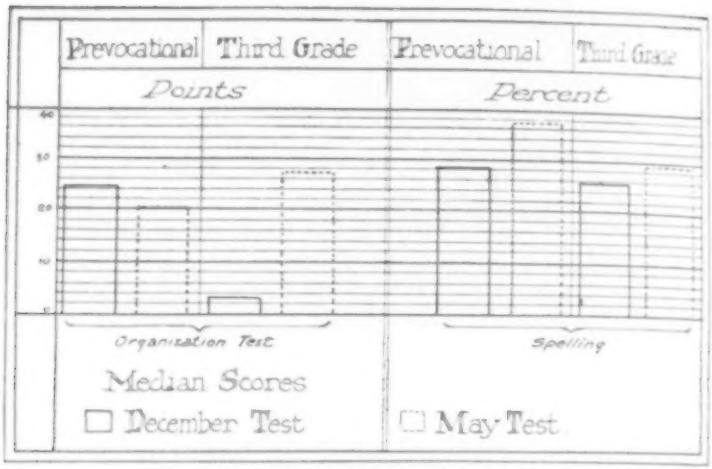

FIGURE 4. COMPARISON OF GROWTH OF A PREVOCATIONAL AND A THIRD-GRADE GROUP MEASURED BY AN ORGANIZATION TEST AND A SPELLING TEST

groups stood the same as far as gains and ability to solve correctly were concerned. In the organization test the prevocational group lost four points while the other gained twenty-four points. The amount of gain here for the third-graders is possibly due to a better understanding of the test after having taken it once, and the last test is probably a better measurement of their abilities than the first one. In spelling the prevocational class seemed to have profited more by their course of instruction over the six-months period, having gained about 9 percent, while the others gained but three percent. However, this is about 7 percent less than the gain made in spelling the previous year. The present test is believed to be a better criterion of their spelling ability, because it consisted of different words, thus avoiding the possibility of any carry over; while in the test of the previous year identical words were used.

Figures 5 and 6 show the scores made in the silent-reading test. Here, too, we observe an equal ability, but a greater gain made by the third-grade group. In the initial test the prevocational boys read on the average 143 words per minute, and 147 words in the final test; the third grade gained from 105 to 145 words, a gain of 44 words per minute, and a rate which exceeded that of the

other group by two words. In ability to interpret the printed page the prevocational boys made a gain of 5.5 percent, the third-graders showed a gain of 17 percent and ranked but 1.5 less than the prevocational group in interpretative ability.

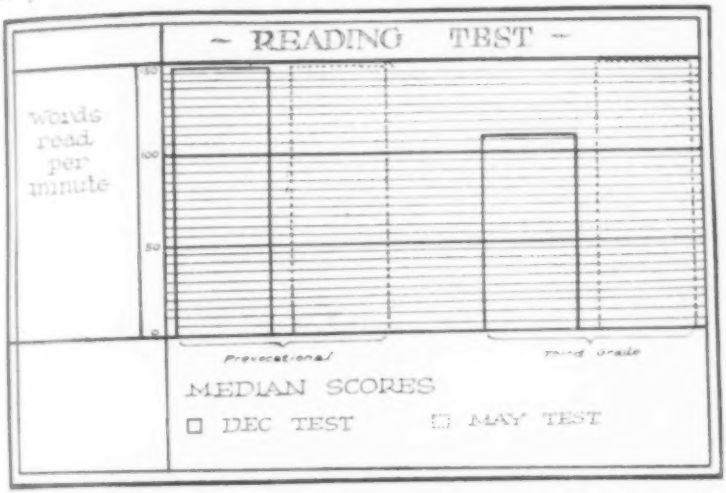

FIGURE 5. COMPARISON OF GROWTH IN RATE OF READING OF A PREVOCATIONAL AND A THIRD-GRADE GROUP

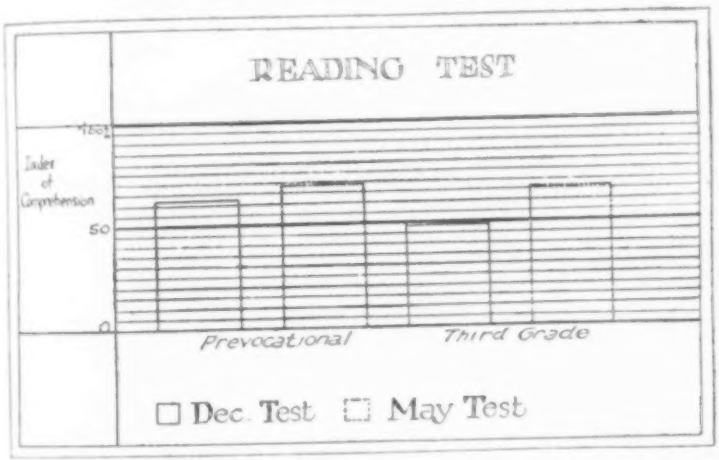

FIGURE 6. COMPARISON OF GROWTH IN READING COMPREHENSION OF A PREVOCATIONAL AND A THIRD-GRADE GROUP

It is quite apparent from the scores made in these tests that the prevocational group has about the average ability of B-third-graders. As a matter of fact, in most of these tests scarcely 20 percent of these boys made scores in excess of the third-grade level, and they did not profit by their course of instruction over this six months period as did the third-grade pupils.

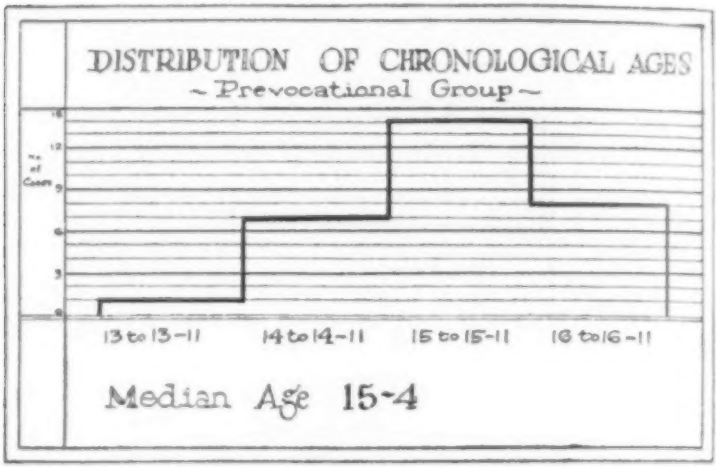

FIGURE 7. DISTRIBUTION OF CHRONOLOGICAL AGES OF THE PREVOCATIONAL GROUP UNDER DISCUSSION

Before drawing further conclusions it might be well to turn from this phase of the experiment and get the status of this particular group of boys from another angle. Figure 7 represents the distribution of chronological ages on June 1, 1920. The boys range approximately from fourteen to seventeen years of age, with fifteen years four months as the median. They are, therefore, judging them by the scores made in the tests, on the average, six years behind in school work.

Figure 8 represents the distribution of intelligence quotients. The range here is from fifty-one to seventy-five with sixty-four as the median. Seventy-one percent of these boys had an I. Q. of less than seventy. In view of this fact it is not at all difficult to account for the low scores and little progress made in the academic branches.

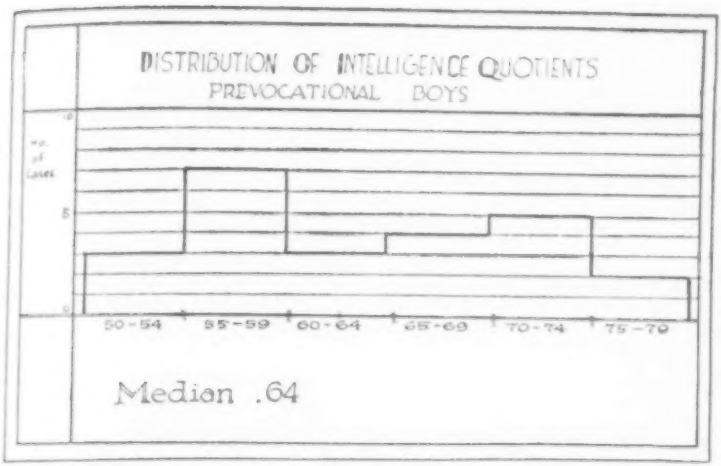

FIGURE 8. DISTRIBUTION OF I. Q.'S OF PREVOCATIONAL BOYS

Figure 9 shows that these boys spent from one to six years, with an average of two years three months, in the special room. This is a sufficient period for them to have profited by both the academic and manual instruction if it were possible for them to do so.

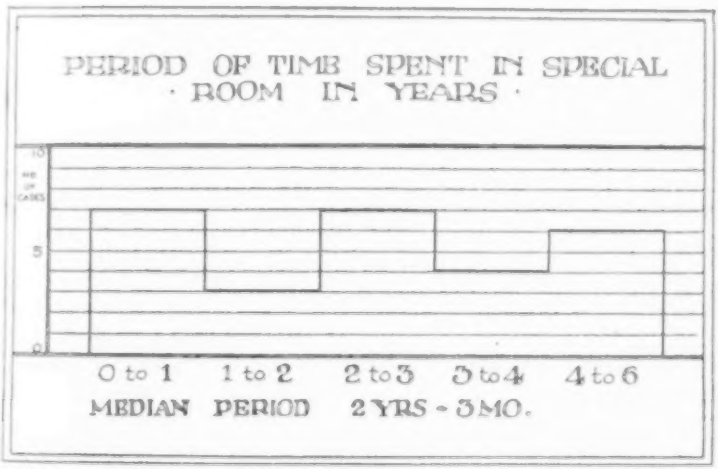

FIGURE 9. DISTRIBUTION OF CASES IN TERMS OF LENGTH OF TIME SPENT IN SPECIAL (PREVOCATIONAL) ROOM

The purpose of the second phase of this investigation was to take an invoice of all the boys who had at some time been members of the Russell Prevocational Class. A list of 125 boys was accordingly made, all of whom had been away from school from one to three years; but, owing to the migratory tendencies of the poorer classes which these boys invariably represent, only 75 were located. Of this number, 27 were found to be unemployed, and 13 of these had not been employed since leaving school while the rest were out of jobs for various reasons. Many had been supplanted by returned soldiers, and possibly many would never have been employed had it not been for the great labor shortage during the war period. Two boys were found in free hospitals, two in the State Institution for Feeble Minded, one in the army, and one in the navy. Two boys had served time in the State Industrial

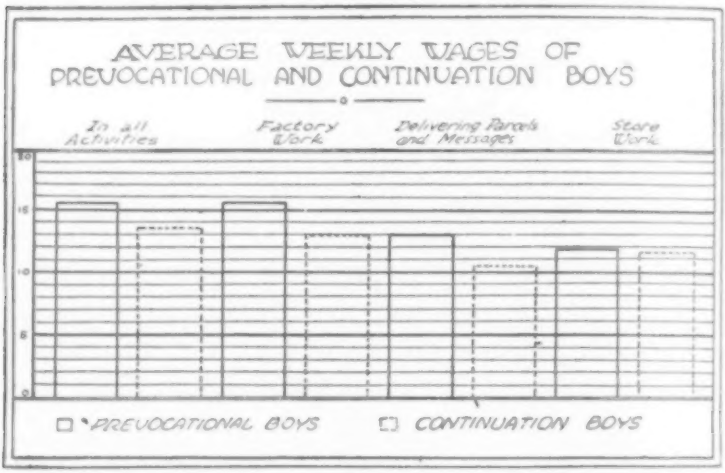

FIGURE 10. COMPARISON OF AVERAGE WEEKLY WAGES OF BOYS ONE TO TWO YEARS OUT OF PREVOCATIONAL SCHOOL AND AN EQUAL NUMBER OF BOYS IN CONTINUATION SCHOOL

School but had been released. Fortunately only one individual had taken upon himself marital duties. He, finding them too strenuous, however, had returned to his parental domicile. Only seven boys were located who had held the same jobs since leaving school. The character of employment and the wages received by these boys seemed to be of vital importance; and in order to make a com-

parative analysis, the average weekly wages of an equal number of boys from the continuation department of the Cass School were checked. Figure 10 shows the results of this part of the investigation. The reader should bear in mind when looking at this graph that all of the prevocational boys were from seventeen to twenty years of age, while the continuation group were all in their sixteenth year. The average wages of the boys from the prevocational class were two dollars per week more than the wages of the continuation boys. A similar situation was found even in the same field of activity as is shown on the right of the figure. In factory work their average excess was two and a half dollars per week; in delivering parcels and messages it was two and a half dollars per week; and in store work twenty-five cents per week.

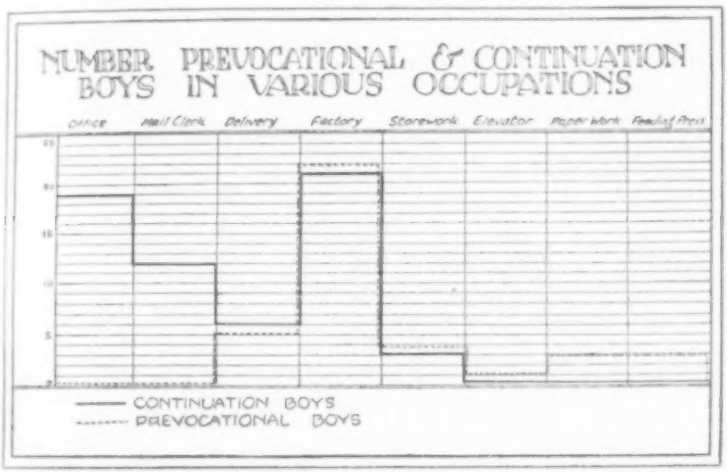

FIGURE 11. DISTRIBUTION OF CASES OF PREVOCATIONAL AND CONTINUATION SCHOOL BOYS BY OCCUPATION

Figure 11 shows the number of boys from both groups employed in various lines of activity throughout the city. The supremacy of the continuation group is here clearly shown, as many of them are found employed in activities for which the prevocational boys are not qualified. It might be added here that of the 75 continuation boys investigated, 57 held the same position throughout the year, and that the eighteen who changed their places of employment apparently did so to better themselves. They were all enrolled in continuation classes, pursuing courses

that had a bearing, in most cases, upon the particular field in which they were employed. On the contrary the prevocational boys seemed to be victims of circumstance with no definite view or purpose in life, getting a job here and there wherever they could find someone to hire them. They seem to be most successful in the factory. No doubt it furnishes the best opportunity for this type of boy because of the better wages and the greater field for stereotyped work requiring a low order of intelligence and initiative. In the factory some of the boys were found to be operating machines, but most of them were doing roustabout work, such as pushing trucks and sweeping floors. No boys were found to be doing wood work or to be making any direct application of the training they had received in the manual training room.

It might be held that an investigation of this kind proves nothing because a repetition a week later would find some of these idle ones employed. However, from the general trend of affairs one would be more likely to find some of the employed ones idle. As a matter of fact, a survey of the 25 boys who left school during the last year, who were not included in the 75 previously recorded, shows similar results. Thirteen of these boys were found to be working, three of whom were not regularly employed; three were selling papers; and only two had remained on the same job since leaving school. One of the twenty-five was deceased, and eleven were known to be idling on the streets.

The conclusions from the investigation seem to the writer to be of lesser importance than the fact that the work was carried on. The inference drawn may be faulty, but the only hope for a more satisfactory solution of this problem as it now stands lies in continued work along these lines. As far as the present survey went, it would seem to challenge the term "prevocational" as applied to this particular room, because the instruction offered there seemed to have no bearing upon the work of these boys after they left. Whether this particular room is a failure in the educational scheme of Detroit depends entirely upon what the room is expected to do. If it is intended for pupils who are retarded and who are later to be reclaimed for the grade, it falls short of its purpose, as only one boy is on record as having returned to his regular grade. If the purpose of the room is to furnish a relief to the regular grades, by eliminating that type of child who does not get on well there, it does all that is expected of it. It is also true that this accomplishment alone is a great piece of constructive work in education. It

would seem, too, that this is about all that can be hoped for under the present order of things. It is generally conceded by those who are best informed that this type of boy will not go very far forward in academic studies, and the present arrangement emphasizes academic instruction. The results of the tests in this experiment are evidences that these boys have not gone very far in academic knowledge, and the I. Q. ratings mean that they have not the possibilities of improving their intellectual status to any great extent.

If any definite conclusion can be drawn from this investigation at all, it is this: that as far as proper returns for efforts expended are concerned, these boys are able to compete with the usual lad far more satisfactorily in the industrial world than they are in the pursuit of academic knowledge. If that be true, it follows that the program for these children should center around the industrial idea. It is not to be inferred that they should be trained specifically for the trades; a general training centering around their probable field of activity is possibly the best that can be done for them. They will be far happier in life if they are able to read, spell, write, do simple arithmetic, understand healthy living, and have regard for law and authority; but outside of these limits they will no doubt get better returns for time and effort expended if their energies are directed along vocational lines.

Certain advantages for a program along these lines can be seen in a centralization plan. For instance, larger groups could be broken up into smaller like units, a departmental plan of teaching effected, a wider range of manual instruction offered, and a placement scheme adopted whereby these boys could be given jobs through the affiliation of the school with the factory or the shop. This scheme could be defended from two angles; first, the apparent adaptability of the children concerned, and second, the economic condition of the parents whom they represent. It might be added here that the average number of children in the families represented by these boys is five, that the father is in most cases a day laborer and a renter, and that every one of the hundred homes visited, required all or part of the boys' daily earnings for their maintenance and support. Many of the circumstances were pressing. If, therefore, a program of this type could be made an actuality, it seems probable that the proposition could be better defended and that the problem of the exceptional child could be more satisfactorily solved for every one concerned.

THE RELIABILITY OF PREDICTION OF PROPORTIONS ON THE BASIS OF RANDOM SAMPLING

BEN D. WOOD
Teachers College, Columbia University

The layman and the statistician of little experience find it difficult to accept with any satisfactory degree of confidence predictions based on proportions of comparatively small random samplings. For example, if it is observed in a random sampling consisting of one-quarter of all the sixteen-, seventeen-, and eighteen-year-old boys in a given city, that 83.4 percent have the father as guardian, what would be the proportion of the remaining three-quarters of such boys who would similarly have the male parent as guardian? The average layman would not even attempt to guess within 10 percent of the truth, and he would probably laugh if someone should venture that it would be 83.4 percent plus or minus 2 percent or less. Again, if for the above sampling it were observed that, for 6.3 percent of the boys, the second year high school was the last school grade completed, that for 1.4 percent sickness was the (reported) cause for leaving school, that for 9.8 percent $18 was the (reported) beginning weekly wage, and that 2 percent left school at the age of 13 years, the average person would be far from ready to accept these as anything like the proportions that would be observed in the total group.

Many will welcome the evidence afforded by an empirical study which recently came to light in the form of a test case which is none the less valid for having been made somewhat clandestinely by a group of skeptics. On December 3, 1918, the Vocational Bureau of the New York State Military Training Commission received a questionnaire card from each of the 6,468 employed boys sixteen, seventeen, and eighteen years of age in the city of Buffalo. About 275 public school teachers filled out the cards for the boys. The same thing was done in every part of the state; and in order to avoid the tremendous task of handling so many cards, the present director of the bureau, Mr. H. G. Burdge, upon assuming charge, gave orders that in certain units random samplings be taken which were to be studied in lieu of the total number of cards for such units. The group of subordinates in

TABLE I. RESULTS OF RANDOM SAMPLING[1]

Description of Item	Percent of Enrollment		
	25	75	100
	1	2	3
I. Guardian of Boy			
Father	83.4	82.4	82.4
Mother	13.3	14.1	13.9
Uncle	0.6	0.6	0.6
Aunt	0.4	0.2	0.2
Stepfather	0.7	0.9	0.9
Stepmother	0.2	0.1	0.2
Brother	0.5	0.5	0.5
Sister	0.2	0.3	0.4
Headmaster or matron	0	0	0
Grandparents	0	0.1	0.1
Others not related	0.6	0.7	0.7
No answer	0	0.04	0.02
II. Number of Children in Family			
One	6	6.3	6.3
Two	11.3	11.8	11.7
Three	14.8	13.7	13.9
Four	13.6	14.4	14.2
Five	14.3	14.6	14.5
Six	11.9	12.6	12.4
Seven	9.8	10.5	10.3
Eight	8.1	7.2	7.4
Nine	4.2	4.1	4.2
Ten	3.0	2.7	2.8
Eleven or more	2.7	2.0	2.2
No answer	0	0.04	0.03
III. Reason for Leaving School			
Financial	9.1	10.1	9.9
Wanted to work	68.4	69.4	69
Sick	1.4	1.2	1.3
Graduated	12.2	11.	11.4
Miscellaneous	0.6	0.3	0.3
Disliked school	8.3	7.9	8.0

Description of Item	Percent of Enrollment		
	25	75	100
	1	2	3
IV. Age Leaving School			
Ten years or under or no answer	0.8	0.7	0.8
Eleven	0.2	0.1	0.2
Twelve	0.6	0.5	0.5
Thirteen	2.0	1.9	1.9
Fourteen	31.6	30.1	30.4
Fifteen	36.9	37.3	37.1
Sixteen	21.5	23.5	22.9
Seventeen	5.5	5.0	5.2
Eighteen	0.9	0.9	0.9
V. Last Grade Completed			
Fourth grade or under or no answer	2.1	2.2	2.2
Fifth grade	3.2	3.4	3.4
Sixth grade	14.5	13.5	13.8
Seventh grade	19.7	20.3	20.2
Eighth grade	23.7	26.9	26.1
1st yr. H. S.	23.8	20.4	21.2
2nd yr. H. S.	6.3	6.2	6.2
3rd yr. H.S.	1.7	2.2	2.0
4th yr. H. S.	1.8	1.4	1.5
Business school	3.2	3.3	3.3
VI. Beginning Weekly Wage			
$ 3.00	10.1	8.6	8.9
$ 6.00	17.4	18.0	17.9
$ 9.00	13.8	15.1	14.8
$12.00	11.2	10.9	10.9
$15.00	14.5	14.4	14.4
$18.00	9.8	9.4	9.5
$21.00	7.7	7.6	7.6
$24.00	5.6	4.7	4.9
$27.00	2.8	3.6	3.4
More than $27.00	0	0	0
No answer	7.1	7.7	7.6

[1] Column 1 shows the proportions in a random sampling of 25 percent (1,617 cases) of Buffalo, New York boys 16, 17, and 18 years of age who have the characteristic indicated at the left. Column 2 shows the proportions observed in a sampling of 75 percent (4,851 cases) of the boys, and Column 3 shows the proportions observed in a 100 percent sampling of the boys (6,468 cases). This table shows parallel columns of proportions for only six of the twenty-four items available for use in this study.

charge of the Buffalo cards was so skeptical that some of them determined to test, *sub rosa*, the wisdom of Mr. Burdge's economy.

Accordingly, the 6,468 cards were put into strict alphabetical order, and every fourth card was extracted. The extracted cards, thus comprising 25 percent of the total and representing 1,617 cases, were sorted and tabulated with the Hollerith machines, as Mr. Burdge had directed. The remaining cards, comprising 75 percent of the total (4,851 cases), were run through the machines for similar sorting and tabulation. Finally, all cards were thrown together and the total 6,468 cards were put through the machines. The results were placed in parallel columns as in Table I. The agreement illustrated ought to put an end to heresy. It is noteworthy that even in the items involving small numbers of cards, the proportions in the three groups are almost identical, clearly demonstrating the sagacity of Mr. Burdge's judgment in the matter.

The parallel columns of the larger table from which Table I is taken afford material for comparing the theoretical with the observed or empirical reliability of the percentile method with random sampling. The formula for the standard deviation of a proportion is

$$\sigma_p = \sqrt{\frac{pq}{n}} \quad \ldots\ldots\ldots\ldots (A)$$

in which p indicates the proportion having the characteristic in question, q the proportion not having it, and n the number of events or cases in which the proportions p and q are observed.[1]

We might test the validity of this formula by comparing the index of reliability which it gives with the index actually observed in the data illustrated in Table I. But the form of the data makes it more convenient to test the formula in a slightly altered form, that is, the form which gives the theoretical standard deviation (σ_{Dp}) of the difference of proportions:

$$\sigma_{Dp} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} \quad \ldots\ldots\ldots\ldots (B)$$

This formula is directly derived from formula (A) by means of the equation for the standard deviation of the difference between

[1] The derivation of this formula is given in detail in chapters 13, 14, and 15 of Yule's *Introduction to the theory of statistics*, C. Griffin & Co., London, 1917. See particularly pp. 257 ff., for all the theoretical considerations of this formula and for the conditions of truly random sampling.

corresponding values of two variables. This equation (Yule, *op. cit.*, p. 210 ff.) may be derived as follows:

Let z be the difference between any two corresponding values, x_1 and x_2 of the variables.

That is, $$z = x_1 - x_2.$$

Squaring both sides of the equation and summing,

$$\Sigma(z^2) = \Sigma(x^2_1) + \Sigma(x^2_2) - 2\Sigma(x_1 x_2)$$

Dividing both members by n,

$$\frac{\Sigma(z^2)}{n} = \frac{\Sigma(x^2_1)}{n} + \frac{\Sigma(x^2_2)}{n} - \frac{2\Sigma(x_1 x_2)}{n}$$

That is, if r is the correlation between x_1 and x_2 and σ^2, σ_1, and σ_2 are the respective standard deviations,

$$\sigma^2_z = \sigma^2_1 + \sigma^2_2 - 2r\sigma_1\sigma_2$$

Assuming that $r = 0$,

$$\sigma^2_z = \sigma^2_1 + \sigma^2_2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Now, since the variabl are proportions,

$\sigma^2_z = \sigma^2_{Dp}$, where Dp stands for the difference of proportions,

$\sigma^2_1 = \sigma^2_{p1}$, and

$\sigma^2_2 = \sigma^2_{p2}$

Substituting these values in (1)

$$\sigma^2_{Dp} = \sigma^2_{p1} + \sigma^2_{p2} = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

and $$\sigma_{Dp} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

Formula (B) would inspire us with more confidence if we had empirical proof that it conforms to fact. If, for example, we should calculate σ_{Dp} for various ranges of percentiles by the use of the formula, and if we should find that the standard deviation of observed differences of proportions in these ranges approximated fairly well the theoretical σ_{Dp}, then we should feel more secure in using this formula in working with percents. These calculations for theoretical and observed standard deviations of the differences of proportions have been made roughly with results as shown in Table II.

The values in the column headed "Observed S. D." were obtained by distributing within each percentile range indicated the differences between every pair of proportions in columns 1 and 2 of the larger table from which Table I was taken, and then

TABLE II. STANDARD DEVIATION OF THE DIFFERENCES OF PROPORTIONS EMPIRICALLY AND THEORETICALLY DERIVED

Percentile Range	Number of Observed Differences from Each Percentile Range	Observed S. D.	Theoretical S.D. from Formula B
50-65	13	1.78	1.43
65-75	14	2.15	1.316
75-85	29	1.756	1.149
85-90	22	1.288	0.950
90-94	25	1.259	0.778
94-97	30	0.7865	0.596
97-98.5	17	0.3937	0.426
98.5-99.5	27	0.2816	0.252
99.5-99.8	10	0.1948	0.1675
99.8-99.9	12	0.0913	.....
99.9-99.97	10	0.0946	.....

calculating the standard deviation in the ordinary way except that the mean was assumed to be zero. Thus, in the table of which Table I is a sample there were 13 pairs of proportions within the percentile range 50–65. The differences between these 13 pairs of proportions gave a distribution with a standard deviation of 1.78 (without correction for assumed $M = 0$).

The theoretical values of the standard deviation in the last column of Table II were obtained by means of Formula B described above, namely $\sigma_{Dp} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$, in which p_1 equals a given proportion in Column 1 of Table I, q_1 equals $(1-p_1)$, and n_1 equals the number of cases (1,617) involved in Column 1 of Table I; p_2, q_2, and n_2 $(=4851)$ are the corresponding values for Column 2 of Table I. For the purpose of this rough verification of the formula, the midpoint of each percentile range was taken as the value of both p_1 and p_2. Thus, for the range 50–65 (midpoint = 57.5), the theoretical standard deviation of the differences of proportions is:

$$\sigma_{Dp} = \sqrt{\frac{(0.575)\,(0.425)}{1617} + \frac{(0.575)\,(0.425)}{4851}} = 1.43$$

It will be noted that the observed standard deviations are larger than the theoretical; but the differences are consistent,

and from a practical viewpoint not large. The differences are due partly to the roughness of the calculations in both columns, partly to the slight inaccuracies involved in carrying the original proportions to one decimal place only, partly to the slight error introduced by assuming that the mean is zero in calculating the observed standard deviations, and largely to the fact that 275 relatively untrained teachers made out the cards. Such considerations as these would justify reducing the denominator in the formula $\sigma_{Dp} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$ quite considerably, so as to increase the theoretical standard deviation systematically. Another influence which makes for a consistent difference in favor of the observed standard deviations is the inadvertent weighting of various differences of proportions by repetitions of sortings involving practically the same or dependent elements. This is notably the case in the second observed value ($=2.15$). This vitiation crept in before the fact of correlated sortings was noticed.

On the whole, the roughness of these calculations does not hide the very strong and unequivocal support afforded by empirical facts for the theoretical reliability of the percentile method with truly random sampling.

A SERIES OF STANDARDIZED DIAGNOSTIC TESTS FOR THE FUNDAMENTALS OF ELEMENTARY ALGEBRA[1]

HARL R. DOUGLASS
University of Oregon

Standardized tests of progress in the study of mathematics may be of two sorts, those which are primarily tests of power and those which are also tests of rate of work. Mathematical tests and scales, however, have been mostly of the latter type and are given with time limits such that few, if any, pupils are able to do all of the exercises. On these tests the pupil's performance has been described in terms of the number of exercises done correctly and of the number of exercises attempted in a given time. Such tests, with minor exceptions, consist of exercises of approximately equal difficulty, and the tests have been scored on that assumption. For this reason the exercises selected for each test have been of the same type, testing the same ability, and requiring the same degree of power of solution.

The series of standardized diagnostic tests for measuring power in the fundamentals of elementary algebra described in this article are based on somewhat different principles of construction. The intended function of these tests is to measure the power to do certain types of exercises in elementary algebra. In the out-of-school situations in which algebra will be applied, rate of work is clearly secondary to accuracy and within limits negligible. A pupil's performance on these tests is to be described in terms of weighted values of the exercises constituting the test, and within certain large limits to be mentioned later, with no reference to time. The exercises of the tests were selected so as to test proficiency in a variety of subtypes of operations under each of the fundamental processes of elementary algebra, and so as to afford considerable range in degree of difficulty. *The tests emphasize a more complete measurement of power and a more thorough and minute opportunity for diagnosis at the expense of measurement for rate of work.*

[1] These tests may be obtained from the University Co-operative Store, University of Oregon, Eugene, Oregon. Price, $1.60 per hundred. Those who use the tests under standard conditions should report median scores and number of cases to the author at Eugene, Oregon.

The fundamentals of algebra.—The first step in the construction of a standardized test for measuring school achievement is the definition of what is to be measured, i.e., in what field of learning or in what portion of a selected field ability is to be measured. In order to answer this question for a test in elementary algebra a questionnaire was prepared that requested those to whom it was addressed to designate the processes of algebra, as ordinarily taught in the first year of secondary schools, which they considered fundamental in the sense that addition, subtraction, multiplication, and division are considered to constitute the fundamental processes of arithmetic. This questionnaire was sent to one hundred members of the American Mathematical Association, whose names were selected so as to include an approximately equal division of teachers in secondary and higher schools, to provide for a representative geographical distribution, and to include those known by the then secretary of the association to be interested in the measurement of mathematical proficiency by standardized instruments.

TABLE I. GEOGRAPHICAL COMPOSITON OF THOSE MAKING REPLIES TO QUESTIONNAIRE TO DETERMINE ELEMENTS OF FIRST-YEAR ALGEBRA

Geographical Location	Number of Teachers		
	Colleges and Universities	Secondary Schools	Total
New England	3	6	9
Atlantic States	6	6	12
Southern States	2	2	4
North Central States	9	12	21
South Central States	2	2	4
Rocky Mountain States	2	2	4
Pacific Coast States	3	2	5
Totals	27	32	59

Fifty-nine replies were received. The geographical distribution of the persons who replied is shown in Table I. More than one-third are from the North Central states but all sections of the country are represented. Table II gives a sum-

mary of the questionnaire. College and university instructors agree very closely with instructors in secondary schools with reference to the operations of first-year algebra which should be considered fundamental. Four processes, (1) collections of terms, (2) division, (3) multiplication, and (4) solution of simple equations received almost a unanimous vote. Four other processes received a majority vote.

TABLE II. FREQUENCY OF REPLIES TO QUESTIONNAIRE TO DETERMINE FUNDAMENTAL PROCESSES OF FIRST-YEAR ALGEBRA

Fundamental Processes	Number of Teachers		
	Colleges and Universities	Secondary Schools	Total
Collection of terms (Addition and subtraction)	27	32	59
Division	27	31	58
Multiplication	27	31	58
Solution of simple equations	26	28	54
Solution of simultaneous linear equations	19	22	41
Factoring of type forms	19	21	40
Solution of simple quadratics	18	20	38
Graphing	17	17	34
Transposition	13	16	29
Exponential manipulation	11	14	25
Expansion of binomials	12	13	25
Clearing of fractions	11	14	25
Radicals	12	12	24
Ratio and proportion	10	9	19
Evaluation of formulæ	3	3	6
Forming equations	4	1	5

The content of the tests.—For each of the four processes which were most frequently considered fundamental, there was constructed a test consisting of ten exercises. This number was considered an adequate compromise between the desirability of including a large number of exercises for minute measurement and the desirability of a brief test which would not require an undue amount of time for administration. The following considerations in the selection of the exercises were observed:

1. The exercises selected should clearly require proficiency in the fundamental process for which the test was being constructed.

2. The list of exercises in each test should provide for testing the chief subtypes of difficulty and teaching units in each fundamental process.

3. The exercises should be so selected that a differentiation of power would be possible on the basis of the degree of difficulty involved.

4. For the purpose of complete measurement and differentiation each test should contain one or more exercises which could be solved by only a small percent of first-year algebra pupils.

These principles were exemplified in the following tests:

STANDARD DIAGNOSTIC TESTS FOR ELEMENTARY ALGEBRA

Test I. Collection of Terms[1]

1. Add
$15m$
$12m$

2. Add
$8ab$
$-7ab$
$-5ab$
$+2ab$

3. Subtract
$12a$
$7a$

4. Subtract
$-4ab$
$11ab$

5. Add
$3a - 8b - 6c$
$5a + 7b - 3c$

6. Subtract
$15a^2 - 8a - 3$
$6a^2 + 7a + 2$

7. Add
$7a - 6a^2 + 10$ and $a^2 + 6a - 7$

8. Subtract
$x + xy - 3y$ from $6x - 3xy + 7y$

9. Collect Terms
$a^2 + 8a + 7a^2 + x + 7 + 9x + 4 + 3a =$

10. Collect Terms
$2x - xy + 3y - 3x + 2xy + 7y + 8 =$

Test II. Multiplication

1. Multiply
$9m$
2

2. Multiply
$4x^2$
$3x^2$

3. Multiply
$5ab$
$-2a^2$

4. Multiply
$+3ay$
$-6a^2y$

5. Multiply
$4a - 3b$
$2ab$

6. Multiply
$4a^2 - 2x$
$-3x^2$

[1] These tests are printed on a four page folder, $8\frac{1}{2}$ x 11. The exercises are arranged so that sufficient space is provided for the pupil's work.

7. Multiply
$8x - 3y$
$4x + 2y$

8. Multiply
$7a^2 - 3x$
$-4ax^2 + 2$

9. Multiply
$3a^2bc + 5b^3cd$
$-4ax^2 + 2$

10. Multiply
$7by^2z - 4a^2cx^2 + 9ab^2y$
$9byz^2 + 8a^2bc^3$

Test III. Division

1. Divide $12a^4$ by $2a^3$.
2. Divide $16x^3$ by $-4x^2$.
3. Divide $-18x^3y$ by $2xy$.
4. Divide $-16x^9 + 28x^6 - 24x^3 + 8x$ by $4x$.
5. Divide $20a^3 - 15a^2 - 5a$ by -5.
6. Divide $14x^3 - 28x^2 + 21x$ by $-7x$.
7. Divide $22a^2b^3 - 16a^3b + 8a^2b^4$ by $-2ab$.
8. Divide $36a^3b + 6a^2b^2 - 12ab^3 + 18ab^4$ by $-6ab$.
9. Divide $6a^3 - 18a - 11a^2 + 20$ by $2a - 5$.
10. Divide $r^3 - 19r + 84 - 6r^2$ by $r - 7$.

Test IV. Solution of Simple Equations

1. Solve for x
$4x = 12$.

2. Solve for y
$5y = 20$.

3. Solve for a
$4a - 3 = 17$.

4. Solve for b
$5b + 6 = 18$.

5. Solve for r
$5r - 7 = 63 - 3r$.

6. Solve for b
$3b - 6 = 14 - b$.

7. Solve for y
$$\frac{2y}{3} - \frac{3y}{4} - \frac{y}{6} = 10.$$

8. Solve for a
$$\frac{6a}{7} + \frac{10a}{14} - a = 4.$$

9. Solve for x
$$\frac{2x}{3} + 8 = \frac{3x}{2} + \frac{x}{6} + 14.$$

10. Solve for x
$$\frac{3x}{5} - \frac{x}{2} + \frac{7x}{10} = \frac{1}{2} + 2x.$$

A large number of recent texts in first-year algebra were examined, and exercises were selected therefrom in conformity with the above requirements. For example, in Test II, Multiplication, Exercise 1 is the simplest type of multiplication that may be called algebra, a positive literal by a positive numerical monomial. Exercise 2, somewhat more difficult, consists in multiplying one positive literal monomial by another, involving exponential manipulation, a phase of multiplication. In Exercises 3 and 4, the multiplier is a negative quantity; hence the multiplication involves a new element, that of the laws of signs for multiplication. In Exercise 5 the process is that of finding the product of a positive binomial and a positive monomial; while in

Exercise 6 the multiplier is a negative monomial. In Exercises 7 and 8, the multiplier is also a binomial, Exercise 8 being more difficult than Exercise 7 because of the complexity of the literal factors. Exercises 9 and 10 are more difficult and were inserted simply to differentiate among the abler students.

Method of administration.—Sufficient blank space is provided on the test sheets for all work necessary for the solution of the exercises. A testing of pupils in all of the subtypes of processes and on the varying degrees of difficulty requires that all pupils have an opportunity to attempt all of the exercises. It is evident, however, that for purposes of economy of time pupils may not be permitted to puzzle over the exercises indefinitely. It also seems that the pupils who work extremely slowly do so because of an unsatisfactory grasp on the methods of solution, finding it necessary to retrace some of the work or to spend undue time puzzling over the proper procedure; and they should, consequently, be penalized Hence, a time limit was set on each test which would permit all workers to do all of the exercises provided they knew how to proceed without undue hesitation and did not attempt an exercise more than once. On the basis of these considerations the following time limits were set: Test I, 5 minutes; Test II, 7 minutes; Test III, 9 minutes; Test IV, 8 minutes

Since the purpose of the tests is not to ascertain how many exercises may be done in a given time, and since the emphasis is placed upon accuracy rather than speed, a measure of proficiency may be secured which is not affected by the element of rate of work, and which more closely resembles the demands of the life situations in which algebraic manipulations in question will be applied. A psychological element of "hurry" is not aroused, and the likelihood of pupils failing to solve exercises because of excitement or undue speed is reduced materially.

The directions to the pupils suggest that they are to work carefully rather than hurriedly. They are told: "We want to see how many of these exercises you can do correctly. You will be given time enough to try all the exercises if you do not waste any time. Work rapidly but do not hurry. Accuracy will count more than rapid work."

Weighting the exercises.—The tests were given to 938 pupils in fourteen high schools in five states of the middle west and far west. The distribution of scores, the medians, 25- and 75-per-

TABLE III. DISTRIBUTION OF PUPILS ACCORDING TO THE NUMBER OF EXERCISES SOLVED CORRECTLY

Number Exercises Solved	Tests			
	I	II	III	IV
0	10	1	4	16
1	12	5	53	13
2	34	10	53	38
3	70	17	62	47
4	56	19	64	104
5	118	74	71	142
6	144	110	103	238
7	152	144	150	139
8	130	241	138	85
9	128	217	99	59
10	84	50	101	57
Total number tested	938	888	898	938
25-percentile	4.44	5.87	3.82	5.10
Median	7.16	8.27	7.26	6.46
75-percentile	8.88	9.21	8.82	7.76
Quartile deviation	2.22	1.67	2.50	1.33

centiles and quartile deviations, from these papers scored without weighting the exercises of the tests, are shown in Table III. Weights for the exercises of the tests were determined by reference to the curve of probability and expressed in terms of median deviation (P.E. or M.D.). A zero point for each test was determined by methods in recognized use.[3] The weights with reference to these zero points are given in Table IV.

Methods of scoring.—In out-of-school situations where algebraic processes are required, the test of proficiency is accuracy. Answers are either right or wrong. No leniency is shown in business dealings to the individual who "has the right method but who has made a small mistake." For this reason and to facilitate scoring, solutions are graded either as right or as wrong. Full credit or no credit is allowed. Answers being correct except for arrangement of terms are not counted wrong. The sum of the weights of the exercises done correctly constitutes the score of the pupil.

[3] Hotz, H. S. *First-year algebra scales.* (Teachers College Contributions to Education, No. 90, 1918), pp. 62–70; also Trabue, M. R. *Completion test language scales.* (Teachers College Contributions to Education No. 77, 1916), pp. 45–60.

TABLE IV. FINAL WEIGHTS IN P. E. OF EACH EXERCISE OF EACH TEST. (TO NEAREST TENTH ONLY)

Exercise	Tests			
	I	II	III	IV
1	1.4	1.7	1.1	2.0
2	1.6	2.7	2.3	1.8
3	1.8	2.9	2.7	3.1
4	3.3	3.2	2.6	4.0
5	2.0	3.6	2.9	4.5
6	2.8	3.0	2.9	3.6
7	2.6	3.9	3.3	5.9
8	3.1	4.6	3.2	5.7
9	3.9	5.1	3.5	6.0
10	3.5	6.4	3.4	6.0
Total	26.0	37.1	27.9	42.6

Standard scores.—Standard median scores for late in February or early in March taken from 938 papers written by first-year classes in fourteen high schools in five states of the West and Middle West, are as follows: Test I, 16.2; Test II, 24.0; Test III, 16.9; Test IV, 20.0.

Values and limitations of the tests.—These tests possess both certain values and certain limitations; chief among which are:

VALUES:

1. The tests provide adequate opportunity for diagnosis of weakness.
2. Opportunity is provided the classroom teacher to check the effectiveness of teaching with respect to the various types of the four fundamental processes.
4. Accurate measurement of power is possible because of the variety in types of exercises and degrees of difficulty.

LIMITATIONS:

1. But one or two exercises are included for each subtype of operation.
2. Because of the extended time limits, rate of work cannot be adequately measured.
3. The tests apply only to the four fundamentals of elementary algebra as determined by the combined judgments of those replying to the questionnaire.

SCALE OF ATTAINMENT NO. 3.—FOR MEASURING "ESSENTIAL ACHIEVEMENT" IN THE THIRD GRADE

LUELLA W. PRESSEY
Ohio State University

ANALYSIS OF THE CURRICULUM OF THE THIRD GRADE

Reports on two previous "Scales of Attainment" have already appeared.[1] The first of these was an examination for the second grade, the second for the eighth grade. The idea back of all three of these examinations is essentially the same:—the purpose of measuring the *essential achievement* of the grade. Only such subjects as could properly be called "promotion" subjects have been considered, since others, though important from a cultural standpoint, hardly condition a child's school progress. The curriculum of the third grade includes spelling, reading, arithmetic, drawing, singing, and writing. Of these drawing and singing are of little importance from the writer's standpoint since ability in these subjects does not affect the progress of a child through the grades. Also, children are not generally retarded because they cannot write well. The fundamental "promotion" subjects of the grade appear then to be spelling, reading, and arithmetic. And the question, in constructing this examination for the third grade, was as to the most desirable form for tests in these three subjects.

Words for the spelling test could, of course, be obtained from the Ayres' scale. The troublesome problem was whether to have list spelling, a timed sentence spelling test, or some other special form. The construction of a test in silent reading was a more difficult matter. Under the direction of the writer, a special study was made by the members of a large extension class. It was concluded that an ability to read with sufficient understanding to grasp story values was the most important factor for this test. Children in the third grade should be masters of the technic of reading to such a degree as to be able to pay attention to the ideas presented, provided the material is reasonably simple. They

[1] Pressey, L. W. "Scale of attainment No. 1.—An examination of achievement in the second grade," *Journal of Educational Research*, 2: 572-81, September, 1920. Pressey, S. L. "Scale of attainment No. 2.—An examination for measurement in history, arithmetic, and English in the eighth grade," *Journal of Educational Research*, 3: 359-69, May 1921, complete.

should be able to read short stories with understanding and with a certain grasp of the story as a whole. The point most emphasized was that the reading matter selected for the test should be coherent. Isolated sentences, bits of incidents, short selections without any particular story value, would not give measures of greatest use to the teacher, since they would not demand of the child that he grasp the story.

Measurement in arithmetic remained to be considered. In the third grade there is intensive drill in the fundamental operations. But there is more than this; a real beginning is made in solving problems. The children should learn to read a simple problem, interpret it for themselves, and apply the needed process. So it appeared that both skill in the fundamental operations and ability to solve simple problems should be involved in any test of third-grade work in arithmetic. Though the proportion of problem work to drill in fundamentals varies a great deal from one teacher or from one school to another, all teachers and all schools seem to have some work in both fields.

It was decided that the test in spelling should involve the spelling of words in their proper setting in a sentence, that the reading test should require the grasping of the meaning of a coherent story, and that the arithmetic test should involve both the fundamental operations and simple problems. Upon the basis of this analysis the scale was built.

The Scale and Its Construction

The scale appears on a four page folder, each page 9 by 6 inches. The spelling test—with the lines for name, age, etc.—is on the first page. It is so placed because the children cannot know what to do on this test until told by the teacher and because study of the page in advance by those first receiving their papers is therefore of no advantage. The arithmetic test is on the back page and is given immediately after the spelling. The reading test occupies both of the middle sheets, the items being so arranged that the one at the end of the first page is continued on the second, thus starting the children on the second sheet almost without their knowing it.

Directions for the test appear on a single sheet of the same size. All the directions, except those for the spelling test, are of the "question-and-answer" type, since this method has proved

more successful with young children than anything more formal.[1] The spelling test requires about five minutes to give; the arithmetic test, seven minutes; and the reading test, eight minutes. The children thus work twenty minutes; and the time needed for giving directions, passing out blanks, etc., makes up a total of about half an hour for the entire examination.

Scoring of the tests is reasonably simple. The spelling test takes only about half a minute to rate; the arithmetic test requires about the same amount of time, and the reading test no longer. In all, it is possible to rate the test and obtain the total score in about two minutes. A teacher can thus rate the blanks for a class of thirty-five in not much more than an hour. The total score on the examination consists simply of the sum of the scores on the three tests.

The Spelling Test

In regard to form, the spelling test is somewhat unusual. It consists of sentences such as the following:

1. My ———— told me to go.
2. It is very cold in ————.
3. She has a great deal of ————.

The teacher explains to the children that each sentence on the page has a word left out, and that she will tell them what word they are to write in. She then reads the sentence as it should be, tells which word is to be written, and then repeats the whole sentence. Thus, the directions for the three sentences given above are:

> Look at the first sentence. It should be "My mother told me to go." "Mother" is the word you should write on the line where it has been left out. "My mother told me to go."
>
> The next sentence should be "It is very cold in winter." "Winter" is the word you should write. "It is very cold in winter."
>
> The next sentence should be "She has a great deal of money." "Money" is the word you should write. "She has a great deal of money."

By giving the word in its context, the child's attention is not entirely centered upon the spelling of the word. Nothing has been said about a spelling test, nor is there anything on the page to indicate to the children that they are being tested specifically in spelling; they are simply to "write in the word that is left out."

[1] Pressey, L. W. "A group scale of intelligence for use in the first three grades." *Journal of Educational Psychology*, 10: 297-308, September, 1919.

This setting of the spelling problem also gives the children the word in its context without making them write an entire sentence. The method would seem to combine certain of the merits of both list and sentence spelling, since the writing of only one word is required, yet the words are given unmistakably in their context. It should be added that the words were selected from the Ayres list, Columns G to Q.[3]

The Arithmetic Test

The arithmetic test represents an attempt to eliminate writing in the solving of problems since it is ability in arithmetic, not speed of writing, that is being measured.[4] The scheme has been adopted of asking a question, followed by four answers. Only one answer is correct, and the child is to make a mark around it. Sample problems from this test are given below:

6. How much is 8 − 3? **6 5 11 24.**
10. How much is 6 x 3? **18 12 9 3.**
13. How much is 13 + 0? **12 11 14 13.**
16. How much is 33 + 6? **9 39 12 33.**
20. If you bought a bottle of ink for 10c and some candy for 4c, how much would you spend in all? **10c 6c 14c 40c.**
24. Doris has 7 books, Frank has 3 times as many. How many has Frank? **14 21 10 4.**

As is evident from the type of question it is expected that the children will do most of the work in their heads. In fact, on the papers of over a hundred children recently tested there were only two or three marks in the margins. The first four questions are used as examples; two of these deal with the fundamental combinations and two with simple problems. The children seem to

[3] The sentences are not arranged in the order of difficulty of the words to be spelled. The reason for this is psychological. If the list is in order of difficulty, the child feels that the words are constantly becoming harder. He senses that the worst is yet to come. Glancing down his paper, he finds that there are still 10 more lines and he knows that if the words keep on getting harder he will shortly be beyond his depth. As a result he becomes discouraged and misses some of the words he might really spell if he were not so apprehensive. It is surely of no particular importance—since the entire list is given in any case—that the words should appear in any special order, provided the *same* order be used for all.

[4] Thorndike, E. L. and Courtis, S. A. "Correction formulæ for addition tests," *Teachers College Record*, 21: 1-24, January, 1920.

find no difficulty with this method of presenting the problems.[5] In fact, they seem able to work without fatigue for a considerably longer time than is possible by the usual method of working out the problems on paper. The saving in time is also of advantage.[6]

It seems reasonable, then, to suppose that the form of the test is satisfactory. It is very certain that the children like it, it does not have a bad psychological effect upon them, as is the case with other arithmetic tests with which the writer is acquainted, and it appears to be completely applicable to this grade.

The Reading Test

The reading test is more nearly similar to other standardized forms than is the case with the spelling and arithmetic tests. It has, however, from a psychological point of view, one rather important difference. It is made up of paragraphs each of which relates a fairly coherent story or incident. Following each paragraph are four questions. The questions are answered in just the same way as with the arithmetic problems; that is, the child selects the correct answer—from four possible ones—by drawing a line around it. Being already familiar with this type of response as exemplified in the test just preceding, the children find no difficulty in understanding what they are to do. A sample passage appears below:

Once a bright star wanted to come down to earth and be a flower so that she might be near little children who would love her. First, she tried being a white rose, but the children were afraid of the thorns. Then she tried being a daisy, but the children didn't see her because she was so small. Then she

[5] It will be noticed that the wrong numbers are those which will be obtained if the wrong process is used. Thus, in the 20th item presented above, if the child subtracts or multiplies (he can't divide) instead of adding, he will obtain either 6 or 40 as an answer. Both these numbers appear in the list, so he will not be deterred from his wrong idea for lack of an answer that agrees with it. As far as possible, wrong answers have been anticipated in this way.

[6] In an arithmetic problem test recently devised there are 10 problems; the children (of the third grade) are allowed to work for a half hour on these ten. Certainly such a test takes on the character of a measure of endurance. It also leaves out, to a very great extent, the element of speed which is of importance in doing arithmetic. Readiness in grasping a problem, quickness in making the necessary combinations, are surely as essential to real ability in arithmetic as the power to solve problems provided one is given time,—for in real life (for which the school is presumably preparing) one certainly has to make change and perform other simple mathematical operations within a very short time limit.

went up on the cliffs and became a dew-drop, but the children could not climb so high. Finally, she lighted softly on the waters of a shallow pond and became a beautiful white water-lily. Then the star was happy because she was near little children and they loved her.

1. What did the star like best to be? **rose dew-drop water-lily daisy**.
2. What did the star want to be near? **cliffs pond children earth**.
3. What flower were the children afraid of? **rose daisy pansy lily**.
4. Where did the dew-drop grow? **pond cliff field lake**.

This passage should be compared with similar passages from other reading tests for this grade to appreciate the difference in "story value" between it and those appearing in other tests. The selections given below are from various reading tests that employ a more or less similar method.

A crab who lived in a sand-hill was sitting at his door in the sun eating a rice cake. An ape went by carrying an orange seed.

Where did the crab live? .....

John had two brothers who were both tall. Their names were Will and Fred. John's sister, who was short, was named Mary. John liked Fred better than either of the others. All of these children except Will had red hair. He had brown hair.

1. Was John's sister tall or short? .....
2. How many brothers had John? .....
3. What was his sister's name? .....

This book is lying on the desk (a picture of a book, face open, is presented just above the paragraph), but it is hard to make it stay open. With your pencil draw a single straight line to represent a ruler lying across the book to hold the pages open. Be sure to make the line from one side to the other, across the book, instead of making it go up and down.

It is the writer's contention that the above passages have no story value. They are simply isolated *reading exercises*. If the reading matter of a passage doesn't tell a story or develop an idea, what is the use of being able to read it? Such reading matter exists—fortunately—for the most part, only in tests. In all other reading there is usually a story, a description, or an exposition, of some sort of significance. Again, the question of gaining the interest of the children is a vital one. Surely no child could be really interested in such selections as those given above. They don't start anywhere, and they don't arrive anywhere, and they aren't about anything; indeed, it would take a statistician to identify correctly the family described in the second passage. There is a distinct movement now on foot to give children in school reading matter that there appears to be some interest in

reading; such tests as the above are certainly not in the spirit of this movement. Incidentally, there is considerable doubt as to whether the ability to gain a coherent idea of the passage read is at all tested by these tests. They would seem rather to test a type of mental alertness.

Results and First Norms

The separate tests of the scale were first tried out by teachers in university extension classes. From these results, items were selected and a trial form of the scale printed. This trial form was given as part of a survey of Bedford, Indiana.[7] The results showed the need of certain revisions which were accordingly made in the final form which is now ready for distribution.[8] Since making the final form the following tentative norms have been obtained (October testing):

Norms (October)

	Median	No. Cases
Spelling	12	198
Arithmetic	9	198
Reading	11	198
Total Score	30	198

The writer has done very little in the way of validation of the scale. But the test most in need of validation has perhaps been sufficiently investigated. Several teachers have been asked to make ratings of their pupils as to their ability to read with understanding. Correlations of the reading test results with these ratings vary from +0.60 to +0.82. This would seem to afford considerable evidence that the test is really measuring the ability of the children to comprehend what they read.

About the spelling test, there can be little question. Children in the third grade do little in the way of using spelling as a tool in writing. Accordingly, sentence spelling would hardly be applicable to the grade. The present test is presumably at least as good

[7] The writer wishes to express her indebtedness to Superintendent E. W. Montgomery of Bedford, Indiana, for his cooperation in this and many other problems.

[8] The blanks may be obtained in quantities from the Department of Psychology of Indiana University at $0.90 per 100.

a measure of spelling ability as list spelling—and it avoids some of the objectionable features of the latter. There might be some question about the arithmetic test. There is surely little difference between the material used in the test and that used in the school room; any difficulty would be with the method of indicating the answers. Since no child has yet made a zero score on this test, the method of indication would not seem to be so difficult as to interfere seriously with the solution of the problem.

Certain Fundamental Factors Concerned with Test Building

In conclusion, the writer would like to point out one or two rather fundamental aspects of testing and of test construction in the elementary grades which are exemplified in this scale, but which should appear in all scales for use in these early grades—and perhaps in the upper grades as well.

1. One of these fundamental points is the matter of motivation. The writer feels sure that motivation by interest is the only reasonable, indeed the only possible, way with young children. A test that relies upon school discipline for its motivation cannot but be unsound from a psychological standpoint. A motivation coming from interest vitalizes the test situation as nothing else can, and should be sought after wherever possible.

2. The participation of the teachers themselves in the making of a test, and in the selection of the materials for it, is of great importance. They are intimately in touch with the teaching situation. The test builder sometimes, to judge from his tests, is not. There are at present many scales on the market that are rightly condemned by teachers, because they are not closely in touch with the teaching, or are not adapted to the children, or are not fitted for use by the teachers. In the present instance, the writer has contributed little save the method of presenting the material and the technical part of test construction. The rest is the result of the study and observation of competent teachers in test making. A much greater participation of those who are going to use the tests is to be hoped for.

3. Achievement tests in "battery" form are now becoming popular. In such an instrument, tests in more than one subject are included. Under such circumstances, it is easy for an examination to become elaborate and cumbersome. The manuals of directions

for using the test may become prolonged to thirty or forty pages; the time required to make preparation for the giving and the scoring of the test amounts to hours; and the interpretation of the scores finally obtained becomes so involved as to need a personally conducted tour through the results. It has been the writer's intention to keep clear of such elaborateness in giving, scoring, handling, and interpreting. It is possible to keep such achievement scales simple and easy to use. He who does not do so is giving an example of intellectual indolence and poor workmanship. Elaborate scales can be avoided; but an earnest teacher is likely to be led into the use of them if she does not first stop to reckon up the amount of time she must invest in the performance.

Summary

The paper presents a brief scale of attainment for measuring the essential progress in the third grade.

1. Tests in reading, spelling, and arithmetic are included. The form of these tests is, in some respects, new.

2. Validation and first norms are presented.

3. Three suggestions are made concerning fundamental aspects of testing, especially in the early grades: (a) that the test motivation come from interest rather than school discipline, (b) that teachers should be allowed to participate in the building of tests, and (c) that achievement scales should, and can, be kept sufficiently simple in construction to be of great use to teachers.

AN ANALYSIS OF THE CONTENT OF SIX THIRD-GRADE ARITHMETICS

F. T. SPAULDING
Scarborough School, Scarborough, New York

Recent investigations of the content of the elementary-school course of study have centered attention on the utility and effectiveness of the subject-matter which the course of study includes. The Committee on Minimum Essentials and the Committee on Economy of Time in Education, as well as numerous individual investigators, have proposed more or less definite standards for each of the elementary-school subjects. The present study of six third-grade arithmetics was intended to determine the extent to which certain well-known primary texts approach these standards.

Specifically stated, the purpose of this study was twofold: first, to determine the exact nature of the arithmetical work presented; and second, to provide a basis for a judgment of the extent to which the textbooks studied make an appeal (a) to the immediate needs and interests and (b) to the probable future needs and interests of the pupils using them.

In selecting the textbooks for study an effort was made to choose those which are in wide use at the present time or which have been widely used in the past. The six books selected represent a period of fourteen years in the development of arithmetic texts. Listed according to recency of publication they are:

Stone, J. C., and Millis, J. F.: *New Stone-Millis Arithmetic.* (Primary.) Benj. H. Sanborn & Co., 1920.

Chadsey, C. E., and Smith, J. H.: *Efficiency Arithmetic.* (Primary.) Atkinson, Mentzer & Co., 1917.

Hoyt, F. S., and Peet, H. E.: *Everyday Arithmetic.* (Book I.) Houghton Mifflin Co., 1915.

Walsh, J. H., and Suzzallo, H.: *Walsh-Suzzallo Arithmetics.* (Third Year.) D. C. Heath & Co., 1914.

Wentworth, G., and Smith, D. E.: *Arithmetic.* (Book I.) Ginn & Co., 1911.

Milne, W. J.: *Progressive Arithmetic.* (First Book.) American Book Co., 1906.

The method pursued was substantially that used by Wise[1] and Monroe[2] in their studies of arithmetic problems. The third-grade material presented in each book was divided into two classes—(1) examples, i.e., drill and test work, in which the operations to be performed were indicated for the pupil; and (2) problems, or work in which the method of solution was not directly indicated. To make the results as nearly comparable as possible, each separate example or problem was counted as a unit, even though the author had included several such in a single task. Tables of numbers, from which the teacher or the pupil could form an almost limitless series of examples, were omitted in the classification.

Examples and problems were then classified separately as to the arithmetical operations or combinations of operations involved in each, including the use of fractions. Where fractional numbers were employed merely to indicate division (as ½, ¼, etc.) the examples and problems were classified under "Division," classification as "Fractions" being restricted to examples and problems using fractions as such.

Problems were further classified (1) as to general subject-matter, according to a modification of the scheme adopted by Monroe, and (2) as to the use of measurements and the types of measurements employed.

The totals of examples and problems found in the six books, with the proportions of each, are presented in Table I. It is noteworthy that the totals vary from 1777 (Chadsey-Smith) to 3106 (Wentworth-Smith)—a range of nearly 1400. There has been little standardization, apparently, of the amount of material to be covered in this grade. Wide variation is also evident in the proportions of examples and problems presented. There would seem to be a tendency, however, toward a greater proportion of problems in the more recent books—a hopeful sign (assuming that the problems are of the right type), in view of the present emphasis on the need for making schoolwork vital and concrete, rather than abstract.

[1] Wise, C. T., "A survey of arithmetical problems arising in various occupations." *Elementary School Journal*, 20: 118-36, October, 1919.

[2] Monroe, W. S., "A preliminary report of an investigation of the economy of time in arithmetic." *Sixteenth Yearbook of the National Society for the Study of Education, Part I*, chapter 7.

TABLE I. NUMBERS AND PROPORTIONS OF PROBLEMS AND EXAMPLES CONTAINED IN SIX THIRD-GRADE ARITHMETICS

Texts	Number			Percent	
	Examples	Problems	Total	Examples	Problems
Stone-Millis (1920)	1,526	751	2277	67.0	33.0
Chadsey-Smith (1917)	1,334	443	1777	75.1	24.9
Hoyt & Peet (1915)	1,887	667	2554	73.9	26.1
Walsh-Suzzallo (1914)	2,513	582	3095	81.2	18.8
Wentworth-Smith (1911)	2,344	762	3106	75.5	24.5
Milne (1906)	2,059	442	2501	82.3	17.7
Average	1,944	608	2552	76.2	23.8

Table II shows the distribution of examples according to the types of operations involved. With the exception of Milne, the books are practically agreed in requiring the use of a single one of the four fundamental operations in the solution of over 90 percent of the examples and of more than 85 percent of the problems—a practice supported by the findings of Wise in his study of problems taken from everyday experience. But in the relative amount of space devoted to each of the four single operations the books vary widely—so widely that the median is of little value as an indication of general practice. Here again practice has apparently been determined in each case by the arbitrary judgment of the textbook maker, rather than by reference to any sort of objective standard.

The variation in proportionate emphasis on the fundamentals between examples and problems within the same book is as worthy

TABLE II. PERCENTS OF EXAMPLES AND PROBLEMS INVOLVING THE VARIOUS OPERATIONS

Combinations	Stone-Millis			Chadsey-Smith			Hoyt & Peet			Walsh-Sozallo			Wentworth-Smith			Milne			Medians		
	Examples	Problems	Total	Examples	Problems	Total	Examples	Problems	Total	Examples	Problems	Total	Examples	Problems	Total	Examples	Problems	Total	Examples	Problems	Total
Addition	26.7	20.5	24.7	22.0	26.5	23.1	44.4	21.3	38.4	22.3	21.1	22.1	32.5	19.7	29.3	12.0	5.4	10.8	24.5	20.8	23.9
Subtraction	15.1	12.9	14.4	13.0	10.8	12.5	29.3	19.8	26.7	13.5	17.7	14.3	27.3	21.5	25.9	10.8	7.2	10.2	14.3	15.3	14.4
Multiplication	28.2	36.6	31.0	25.4	18.5	23.6	10.7	21.4	13.5	29.4	32.0	29.9	13.2	30.6	17.6	21.9	27.0	22.8	23.7	28.8	23.2
Division	28.4	27.6	28.1	36.3	32.5	35.3	15.6	23.8	17.7	29.7	22.7	28.4	18.8	20.5	19.2	33.0	12.7	29.4	29.1	23.3	28.3
Addition and Subtraction	....	0.3	0.1	....	0.5	0.1	....	5.5	1.5	0.1	....	....	....	0.4	0.1	0.4	2.7	0.8	0	0.5	0.1
Addition and Multiplication	1.5	0.1	1.1	....	1.1	0.3	....	4.0	1.0	....	0.5	0.1	5.5	2.5	4.8	3.1	8.1	4.0	0.8	1.8	1.1
Addition and Division	....	0.1	....	....	....	....	....	1.4	0.4	....	0.7	0.1	1.8	0.2	1.4	....	....	....	0	0.4	0.1
Addition, Subtraction & Multiplication	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	.2	....	0	0	0
Addition, Subtraction and Division	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
Addition, Multiplication and Division	....	0.1	....	....	1.8	0.5	....	....	....	....	....	....	....	0.1	....	....	....	....	0	0.1	0
Subtraction and Multiplication	....	0.3	0.1	....	....	....	....	2.1	0.6	....	0.7	0.1	....	....	....	0.7	7.0	1.8	0	0.5	0.1
Subtraction and Division	0.1	....	....	....	....	....	....	0.5	0.2	....	0.3	0.1	....	....	....	....	....	....	0	0	0
Subtraction, Multiplication and Division	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
Multiplication and Division	....	1.5	0.5	....	0.5	0.1	....	0.2	....	....	0.2	....	....	....	....	....	11.1	2.0	0	0.4	0.1
All Operations	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
Fractions	....	....	....	3.3	7.9	4.5	....	....	....	5.0	4.3	4.9	0.9	4.3	1.7	18.1	18.6	18.2	2.1	4.3	3.1

of note as the variation between books. Leaving out of consideration a fairly uniform preponderance of addition examples over addition problems (due partly to the use of addition in demonstrating subtraction and multiplication), the tables show in some cases a greater proportion, in others a less, of examples than of problems. Hoyt & Peet, for instance, devote 29 percent of their examples and only 19 percent of their problems to subtraction; whereas 8 percent more problems than examples are devoted to division. In the *Walsh-Suzzallo Arithmetics* the differences are in the opposite direction. Part of the discrepancy is obviously due to the fact that considerably more problems than examples are distributed in the double and triple-operation groups; but this would provide only for smaller proportions of problems than of examples. Lack of a clear conception of the functions of problems and of examples and of the amount of practice desirable in each is responsible for most of the variation.

A hopeful sign is apparent in the partial or the complete elimination of fractions in the later books. Milne devotes 18.2 percent of his examples and problems to fractions as such; Stone & Millis and Hoyt & Peet postpone the study of fractions (except fractions with unit-numerators, used to indicate simple division) to a later grade. In his study of common problems from everyday life, Wise found the use of fractions to be restricted very largely to that of fractions with simple numerators and denominators—a fact which would argue for much less attention to the more complex fractions than has been given by the older arithmetics.

From the point of view of the textbook maker such investigations as that of Wise must serve to indicate the nature of the material which should be included in a course of study, rather than the relative stress which each phase of such material should receive. For it is obvious that the practice necessary to master an operation of arithmetic may be out of proportion, so far as the time element is concerned, to the degree to which the pupil may expect to use this operation in ordinary life; and yet the fact that it is necessary in ordinary life requires its mastery in school. It is dangerous, therefore, to draw conclusions from such figures as we have at hand as to the exact amount of time which should be devoted to each of the several phases of arithmetic-teaching. In so far as the textbooks studied center their attention on practice in the fundamentals they are in accord with the general

principles established by Wise's study. Their weakness is in pedagogical method, rather than in aim; they show no evidence of established standards as to the amount of practice needed to master the fundamentals.

In the classification of the subject-matter of the problems (Tables III, IV, and V), we find a basis for determining the extent to which the texts meet the present and the future needs and interests of the pupils who use them. Here again a very wide diversity is apparent. Of particular significance is the variation in the proportion of problems which relate to no human activity—such problems as, "How many feet have 16 birds?" or "A book cost $1\frac{3}{4}$. How much more than $1\frac{1}{4}$ is that?" The Hoyt & Peet *Everyday Arithmetic* is the only one which has a clean slate in this respect; the Walsh-Suzzallo text is the worst offender, with 4.1 percent of its problems (24 in all) of this worthless type. The tabulation shows an encouraging tendency, however, toward the elimination of this sort of work from the later books.

There is a much larger group of problems (49.6 percent of those in the Wentworth-Smith book) which give valuable practice to the pupil and are of a sort which he will frequently meet, but which as presented can be identified with no particular activity. Of this type are such problems as, "How many pecks are there in 8 bushels?" or, "What is the perimeter of a lot 180 feet square?" In their lack of appeal to the pupil and their lack of connection with concrete activity, they are of much less value, of course, than problems which possess definite significance. As with the problems relating to no activity, the proportion of such problems seems to be decreasing in the more recently published books.

With the exception of Wentworth-Smith the books are fairly agreed in devoting about half their problems to Home Activities, Personal Activities, and the Activities of Children. Five books out of the six agree in relating about 10 percent of their problems to Home Activities; in the other two classifications just mentioned there is considerable divergence. Within all of these fields the quality of the problems presented counts for so much that without a more exact classification than that of the present study, and without more definite standards of evaluation than have yet been developed, we can draw no worth-while conclusions as to present tendencies and limitations.

For the distribution of problems dealing with occupations we have standards of a sort. If we consider these problems in the

TABLE III. DISTRIBUTION OF PROBLEMS ACCORDING TO SUBJECT-MATTER

Texts	1. Occupations														2. Home Activities		3. Personal Activities		4. Activities of Children		5. Relating to No Activity		6. Incapable of Identification	
	1. All Occupations		A. Agriculture		B. Trade		C. Transportation		D. Public Service		E. Industry		F. Others											
	No.	Per-cent	No.	Per-cent	No.	Per-cent	No.	Per-cent	No.	Per-cent	No.	Per-cent	No.	Per-cent	No.	Per-cent	No.	Per-cent	No.	Per-cent	No.	Per-cent	No.	Per-cent
Stone-Millis	240	31.8	97	12.9	113	15.0	8	1.1	10	1.3	1	0.1	11	1.4	83	11.1	56	7.5	249	33.2	5	0.7	118	15.7
Charlsey-Smith	59	13.3	18	4.1	10	2.3	13	2.9	8	1.8	1	0.2	9	2.0	40	9.0	57	12.9	185	41.8	1	0.2	101	22.8
Hoyt & Peet	195	29.5	41	6.1	81	12.2	29	4.4	27	4.2	10	1.5	7	1.1	75	11.2	151	22.6	191	28.5	0	0	55	8.2
Walsh-Suzzallo	140	24.1	61	10.5	62	10.7	10	1.7	1	0.2	6	1.0	0	0	72	12.4	103	17.7	141	24.2	24	4.1	102	17.5
Wentworth-Smith	74	9.6	17	2.2	46	6.0	4	0.5	4	0.5	0	0	3	0.4	82	10.8	125	16.4	79	10.4	24	3.2	378	49.6
Milne	89	20.2	11	2.5	50	11.3	10	2.3	4	0.9	0	0	14	3.2	77	17.4	51	11.5	80	18.1	7	1.6	138	31.2
Average	132	21.4	41	6.4	60	9.6	12	2.2	9	1.5	3	0.5	7	1.2	72	12.0	91	14.8	154	26.0	10	1.6	149	24.2

light not merely of the pupils' present needs, but of their future interests and of the possibility of vocational and civic enlightenment through the work in arithmetic, we find not only much variation in individual texts, but a considerable difference in treatment between the earlier and the later books. The subject-matter in the problems in the earlier books appears to have been selected almost entirely because of its adaptability to the text-book maker's purposes in providing practice in the fundamentals, rather than because of any consideration of the value of the subject-matter itself. The grouping of problems is for the most part heterogeneous and pointless, so far as their subject-matter is concerned. In the later arithmetics it has been the avowed purpose of the authors, attested not merely by their introductory statements but by the arrangement and selection of problems as well, to provide subject-matter which shall make definite appeal to the needs and interests of the pupils. In these later books we find, therefore, first, a distribution of problems over a wider field of interests and activities, and second, a more thoughtful apportionment of problems to the various occupational groups. The first characteristic becomes apparent when we consider the percents of problems represented in each group. The three last published books afford problems under all six occupational headings, whereas the earlier texts tend seriously to neglect various fields—

TABLE IV. A COMPARISON OF THE DISTRIBUTION OF PROBLEMS ACCORDING TO OCCUPATIONS, WITH THE DISTRIBUTION OF THE WORKING POPULATION OF THE UNITED STATES

	All Occupations	A. Agriculture	B. Trade	C. Transportation	D. Public Service	E. Industry	F. Other
Census of 1910	*100.0*	*33.2*	*9.5*	*6.9*	*1.2*	*27.9*	*21.3*
Stone-Millis	100.0	40.4	47.1	3.3	4.2	0.4	4.6
Chadsey-Smith	100.0	30.5	17.0	22.0	13.6	1.7	15.2
Hoyt & Peet	100.0	21.0	41.5	14.9	13.9	5.1	3.6
Walsh-Suzzallo	100.0	43.6	44.3	7.1	0.7	4.3	
Wentworth-Smith	100.0	23.0	62.2	5.4	5.4		4.0
Milne	100.0	12.4	56.2	11.2	4.5		15.7

TABLE V. DISTRIBUTION OF PROBLEMS INVOLVING MEASUREMENTS

Texts	Problems Dealing with Measurement		Linear		Square		Cubic		Land		Dry		Liquid		Weight		Time	
	No.	Percent of all Prob.	No.	Percent[a]	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent
Stone-Millis	214	29	24	11.3	8	3.4	..	....	..	....	41	19.2	65	30.4	36	16.9	40	18.8
Chadsey-Smith	83	19	26	31.3	..	....	..	....	..	....	15	18.1	13	15.7	7	8.4	22	26.5
Hoyt & Peet	154	23	34	22.1	..	....	..	....	..	....	..	....	12	7.8	9	5.8	99	64.3
Walsh-Suzzallo	139	24	29	20.8	2	1.5	..	....	..	....	18	13.0	20	14.4	26	18.7	44	31.6
Wentworth-Smith	168	22	30	17.9	28	16.7	..	....	..	....	27	16.1	24	14.3	22	13.0	37	22.0
Milne	116	26.2	20	17.2	9	7.8	22	19.0	..	....	..	....	15	12.9	9	7.8	41	35.3
Median	147	24	28	19.4	5	2.5	0	0	..	....	17	14.6	18	14.4	16	10.7	41	29.1

[a] Percent of measurement problems.

notably Agriculture and Industry (Wentworth-Smith and Milne), Transportation (Wentworth-Smith), and Public Service (Walsh-Suzzallo and Wentworth-Smith). The second point is confirmed by an inspection of Table IV, in which the percents of occupational problems belonging to each group are presented for comparison with the distribution of workers by vocations according to the census of 1910 (adapted from Monroe). Though the distribution of problems diverges to a considerable extent from that of the industrial population, there is yet evident in the later books some degree of approach to this standard distribution. The over-emphasis on trade is not a serious defect, since this is an occupation with which every adult (and nearly every child) has more or less to do. The most serious neglect appears in the field of industry, which, though claiming nearly 28 percent of the nation's workers, is represented at most by but 5 percent of the problems.

Our final analysis is concerned with the number of measurement problems and with the types of measurement involved (Table V). With but one exception (Chadsey-Smith), the six textbooks are fairly agreed in devoting about one-fourth of their problems to some type of measurement. Beyond this point agreement ceases. There is a tendency in the later books to postpone study of certain forms of measurement (notably square and cubic measure), taken up in detail in the earlier published texts; but in the amount of space devoted to the remaining forms of measurement (linear, dry, liquid, weight, and time) the books are very widely at variance. There is here very evident need of standardization of the course of study on the basis both of pedagogical and of utilitarian values.

Summary

The very limited number of texts involved in the present study makes the drawing of general conclusions a dangerous proceeding. Until such time as a more extensive investigation is possible, however, the following conclusions are advanced for what they are worth.

1. There is much need for standardization of third-grade arithmetic texts: (a) as to the nature and amount of material presented for study in the course of the year; (b) as to the emphasis placed on each of the fundamental operations; and (c) as to the subject-matter of the problems presented for solution.

2. Present practice in textbook writing apparently tends toward a concentration of attention in the third grade on the fundamental operations of arithmetic, with a postponement of the study of fractions to the upper grades—a tendency the value of which is supported by the findings of Wise and of others with respect to the utility of these phases of arithmetic in practical experience.

3. There is evident also a tendency toward greater emphasis on problem-solving, in contrast to the simple "doing examples" of the earlier books.

4. A study of the subject-matter of problems shows an increasing elimination of those which relate to no human activity, or which can be identified with no activity—and, as a corollary, an increase in the proportion of worth-while, intelligible problems.

5. Textbook makers are apparently coming to appreciate the need for making their problems representative of the fields of activity in which pupils are likely to be engaged.

6. Movement in all these directions toward better third-grade teaching material is as yet, however, ill defined and irregular, resting on no firm basis of educational theory. The pressing need of the present time is for a pedagogically sound definition of arithmetic material (a) in terms of the amount needed to accomplish most economically the desired results, and (b) in terms of subject-matter looking not alone to efficient mastery of the fundamentals but to the proper development of the whole child. Until we have such a definition, the making of textbooks in arithmetic must be (as it has been in the past) guesswork, pure and simple.

Editorials

FINANCIAL RESEARCH

There has been great progress in the administration of our city school systems during the past ten years due to the introduction of more adequate child accounting. In every progressive school system data with respect to retardation and elimination, together with the measurements of the achievements of children are being used as the basis for revision of courses of study and the development of a more adequate organization of schools.

One may question, however, whether any comparable advance has been made in the fiscal administration of schools during the same period. As one reads even the more recent legislation with respect to the distribution of school funds one cannot but be impressed with the unscientific methods which are employed in most of our states. If we accept the ideal of equality of opportunity, we certainly have not yet been willing to fight for its realization in terms of an equitable development of grants-in-aid.

In our city school system there is still need for much more adequate accounting than is commonly found. In most American cities the preparation of the budget and the actual fiscal administration of the school system are less adequate than that commonly found in successful business organizations. If our system of financial accounting permitted the comparison of costs among the several school units making up our city school systems, and if we were able, as well, to report to our public the cost of teaching English, arithmetic, the social sciences, or the cost of kindergartens, of health service, the household arts, and the like, our position in these days of retrenchment would be much more secure.

There is likewise need for a most careful scrutiny of the methods employed in financing building programs throughout the United States. Large expenditures are being undertaken in some cases without an adequate study of increases and shifts in population within the area to be served; and in some cases extravagances have been permitted in the construction of school buildings, which could not possibly be justified were the program of capital expendi-

tures for a period of years put clearly before the boards of education and the public.

We need in state, county, and city school systems directors of research whose main purpose will be the study of the fiscal problem. One can hardly expect that those who are primarily concerned with child accounting, organization, the achievements of pupils, and methods of learning can find the time, nor that they will be equally expert in dealing with the problems of fiscal administration in our school system. It will be equally futile to hope for thorough-going scientific work in this field from bookkeepers and purchasing agents without scientific training. No phase of research that has been undertaken promises more certain returns in increased efficiency and in economy in the administration of public education.

G. D. S.

A GENTLE SUGGESTION

In any plan for training teachers in service, acquaintance with the best in professional writing should be included. A recognition of this is afforded by the development of the reading circle idea. Certain books, adopted for inclusion in the reading circle list, have undoubtedly been widely read. Presumably, they have also influenced practice.

It is well known, however, that a great many of the best suggestions for teachers are found in periodicals rather than in books. To go no further than the pages of our own journal, Dalman's concrete example in the January, 1920 number of differentiated requirements for pupils of varying abilities remains the best answer with which we are familiar to a question which thousands of teachers are asking, namely, "After grouping children on the basis of ability, how shall I go about making a course of study for each group?" Nevertheless, it is clear that Dalman's article, though a great little idea, can't furnish forth a book. Accordingly, it can't be included in the reading circle material.

Again, we know of no more satisfactory analysis of fractions than Kallom offered in the March, 1920 number. He gave a test for proficiency in each type of addition of fractions and the form on which the record was made. This form, when made out, shows the types of examples in which the pupil has difficulty. Teaching and drill on these types follow; and the pupil is then

given another form of the test. He is excused from further work in addition of fractions when he attains a consistently perfect score on one of these tests. Here we have curriculum analysis, diagnosis, individual instruction, and tested results—all organized in a highly satisfactory manner. Yet the teacher must remain unfamiliar with this admirable presentation because it does not appear in a book.

At the risk of being tedious we shall give another instance. Doctor Luella M. Pressey s suggestive article in the September, 1920 number must remain unknown to the vast majority of the teachers who could use it. Yet in this article Doctor Pressey not only adds an important test to those already available (and just when we needed it, too), but she also sets forth a method by which the teacher may construct home-made objective tests of her own. At some future time the book writer—whose service consists for the most part in collecting other people's ideas, and whose only contribution may be his particular juxtaposition of them—will no doubt have something to say about this test. He may even refer to Doctor Pressey's article. But it is certain that his write-up will lack the directness, the force, and the vividness of the original. Moreover, it will be late in arriving upon the scene. If an idea is worth using, it is worth using now.

The worth of these ideas does not depend upon their appearance in a bound book. What Rochester is doing in reading (O'Hern and Hawley), what Boston is doing in geography (Barthelmess), what St. Louis is doing in penmanship (Walker), what the Pacific coast school is doing with intelligence tests (Terman, Proctor, Dickson)—these are among the things which wide-awake teachers want to know and which wide-awake superintendents want their teachers to know.

Why, then, may not subscriptions to periodicals or single numbers of periodicals be included in the material which teachers are encouraged or expected to read? Why this predilection for books? In our judgment there is in any one of a half dozen magazines more timely and stimulating educational writing than can be found in any but the most unusual book.

B. R. B.

Reviews and Abstracts

LEWIS, E. E. *Scales for measuring special types of English composition.*

Mr. Lewis has done a painstaking piece of work and has produced in this monograph several new extensions of the original Hillegas Scale. Judges were carefully selected and trained. Consequently one feels that the values attached to the specimens have a greater degree of reliability than is usually the case. Teachers of English composition will no doubt welcome these scales not only for these reasons but also because of the feeling of ease with which letters can be evaluated in terms of the scale.

While these scales do not meet all the standards set up by the author himself, nevertheless their shortcomings—such as unequal steps between specimens, a single specimen at each step, and lack of distinction between form and content values—are not serious enough to raise a valid objection to their use.

Although no evidence is given to show that letters can be more accurately scored with these scales than with either the Thorndike Extension or the Nassau County Supplement of the Hillegas Scale, it is not improbable that this will be found to be true. On the other hand, had the group of judgments been split into two parts and the scale values found on the basis of the two groups of judgments separately, it would undoubtedly have been evident that the very high coefficients of correlation give an exaggerated impression of the stability of the value assigned to the specimens. In the scales themselves, however, the dropping of the second decimal place tends to correct this impression.

In constructing the scales both the method originally employed by Hillegas and in part the method used by Trabue in deriving the Nassau County Supplement, were tried. A somewhat detailed examination of the specimens of social letters and of narration shows that the two sets of values for the latter differ on the average by 0.26 of a step, or about 1/32 of the length of the scale, while for the former the average difference is only 0.17 of a step For the specimens of social letters the positive and negative differences appear in small groups; in the narration specimens they appear in very large groups. This seems to be due to the fact that the very small differences between the specimens were added together to get the total scale length, whereas in the original scale larger and hence more reliable differences were added together. A very few errors entering in this way seem to account for the larger difference between the two sets of values in the narration specimens.

In view of this condition the values obtained by the shorter method of having the specimens graded is probably more accurate than those obtained by the much more involved process. The work of the author in using the original method has, however, been by no means unprofitable; for, as a result of it, the stability of the original Hillegas values has been further substantiated. Moreover, the method of having specimens graded by competent judges with a scale has definitely established the possibility of obtaining various extensions of the original scale and of securing steps at equal intervals.

M. J. VAN WAGENEN

University of Minnesota

HERTZOG, W. S. *State maintenance for teachers in training.* Baltimore: Warwick & York, Inc., 1921. 144 pp.

The author, at the outset, portrays with desirable emphasis the fundamental importance of relieving the teaching profession of an unfavorable social attitude, and the necessity of affording substantial economic relief to competent members of the profession. This is not new; but until these cardinal defects are remedied the teaching profession will continue to occupy a position of inferiority with respect to other professions and with respect to the industries.

Among the purposes laid down in this volume are the following: (a) To investigate teacher-shortage and to record the best relief measures employed in several states, (b) to investigate the principles, problems, and practices involved in a system of subsidies for prospective teachers as one method of recruiting the profession. These purposes are detailed, (a) by a survey of conditions which reveal an imperative need for financial support to teachers in training, (b) by a review of efforts in foreign countries and in the United States to recruit the profession by means of subsidies to prospective teachers, and (c) by drawing lessons from other professions and the industries in their attempt to win recruits by means of financial assistance.

The author concludes that a policy of subsidies, following the principles and practices which obtain in the recognized professions, would contribute substantially to the relief of the existing teacher-shortage and would in large measure raise the profession out of the slough of social inferiority. The volume merits a careful reading by all those who are vitally concerned with teacher-training and by school administrators who have been far too indulgent in accepting the "finished products" of many of our teacher training agencies.

Your reviewer would have welcomed a discussion of the significance of a teacher-shortage as an opportunity on the part of school administrators to impress the public with the dangers involved in an inadequate supply of well-trained teachers. School administrators appear reluctant to place the responsibility for a shortage of competent teachers where it belongs, namely, on the public. They have lowered standards and made possible the employment of thousands of untrained and incompetent teachers. Report has it that 48 percent of the teachers in one state of our glorious Union have had no training beyond the eighth grade, and that 68 percent have received less than a high-school education. If state laws permit this disgraceful condition, the obvious remedy is the establishment of standards. In states maintaining desirable standards, the practice has been to lower them. Such practice reduces the shortage by opening the gates to incompetents, and accustoms an undiscriminating public to a low standard of teaching.

Nothing could be more wholesome to the profession and more beneficial to the nation than a rigid enforcement of standards entailing consequences for which the public should be held responsible. So long as school administrators are unwilling to place the responsibility on the public, where it rightfully belongs, but are willing to employ incompetents in periods of teacher-shortage, so long will the profession endure economic servitude and occupy a position of social inferiority.

GEORGE F. ARPS

Ohio State University

AVERILL, LAWRENCE AUGUSTUS. *Psychology for normal schools*, (Riverside Textbooks in Education). New York: Houghton Mifflin Company, 1921. 362 pp.

Unfortunately for the training of teachers, mere contact with courses in education, no matter what they are, is in many quarters still supposed miraculously to give professional preparation for teaching. Legal requirements for certification often merely demand a certain number of hours in psychology and education, regardless of whether the courses serve a specific practical purpose or meet specific needs. Similarly, textbooks on education are like the old time physician's shot-gun prescriptions, being labelled as suitable for college or normal school classes or for teachers' reading circles. A textbook that is suitable for groups so diverse in preparation, interests and needs is probably not good for much.

Professor Averill's book has the merit of being designed for a specific group, bearing as it does the title *Psychology for Normal Schools*. It is not a text in educational psychology, but in general psychology, though it includes much of what is ordinarily treated in educational or genetic psychology. On the other hand, it differs very much from the ordinary text in general psychology. The first one hundred and forty-three pages, or twenty-one lessons out of forty-six, are given to a detailed discussion of instinctive and emotional behavior, followed by forty-five pages (six lessons) on heredity. Sensation gets seven pages with little in them but a doubtful complete list of sensations and four pages on the earliest sensations of infancy. Six pages dispose of perception, while the whole subject of the perception of space and time gets twenty lines. Memory is discussed in one lesson of six pages without a word about the laws of learning which are so admirably presented in such a book, for example, as Pillsbury's *Essentials of Psychology*. One lesson likewise disposes of such an important topic as thinking and another of the will and moral development. The remaining lessons deal with the juvenile delinquent, the subnormal child, the gifted child, individual difference, the unstable child, adolescence, and the evolution of the social attitude toward children.

What psychology is of most worth for teachers is a question that will be variously answered. That the type of psychology presented in this book will be accepted for normal schools as better than the conventional type the reviewer doubts very much. The student would complete the course outlined with very little notion of what modern psychology is and without and appreciation of the complexities of mental processes. Nor does it appear that he would be better prepared for teaching by it. Surely to give one lesson to a discussion of the food-getting and hunting responses and only one each to such topics as sensation, memory, and thinking is out of all proportion.

V. A. C. HENMON

University of Wisconsin

STEVENSON, JOHN ALFORD. *Project of teaching*. New York: Macmillan Company, 1921. xvi + 305 pp.

This book is a valuable addition to the growing literature which attempts a systematic examination of the proper definition of the term "Project"; and a constructive, though critical, evaluation of the project method as a practical instrument for school use. Too much of what has been written on this topic either takes the form of propaganda or is so theoretical that it has little stimulation for the average school man or woman. Both of these weaknesses are avoided by Doctor Stevenson. The issues involved are, in general, put clearly and fairly, and are made concrete by considerable

illustrative discussion. The book should be welcome, therefore, as a text even by those who would object to certain of the author's conclusions. Those who have followed most closely the development of educational practice in the last twenty years, will probably raise one or more of the following objections to the treatment of the project as found in this book.

(1) The four standards set up by Stevenson must not be regarded, by implication, as standards for judging teaching in general, but only as standards for judging project teaching. Some will feel, moreover, that these standards do not represent an accurate digest of present and past opinion regarding the project, but rather the author's choice of certain elements among these opinions. Such a selection is, of course, justifiable, but the reader must keep in mind that such a choice has been made.

(2) The book should have contained a section treating the practice of schools which the teaching public has thought of as employing the project method. Certainly there should have been a critical review of the work of Booker T. Washington, at Tuskegee; of Dewey, at Chicago; of Miss Cooke, at Francis Parker; of Bonser, at Speyer School; and of Meriam, at Missouri. Possibly, too, some of the more radical schools—such as that of Miss Johnson at Fairhope—should have been evaluated. In other words, many readers will doubtless feel that there has been too much attention given to those who have written about the project, to the neglect of those who have contributed in practice to the development of the various technics involved.

(3) The various issues involved should have been made clear by a careful review of the chief theories which have been prominent in the last quarter of a century, particularly the theories which have gone under the following names: Herbartianism, Froebelianism, pragmatism, reaction psychology, and the problem method as expounded by Dewey. No doubt the author was led to omit the discussion of all these theories except the last, on account of the difficulty of presenting a clear discussion within the limits of a book of the character of this one.

(4) The distinction between the psychology of thinking and the psychology of learning is not made sufficiently clear. Most of the mistakes which have been made in connection with the problem method and the project method may be traced to this confusion. We learn, to be sure, while thinking, but this is not the only, nor always the most economical method of learning.

Some of these objections are referred to, or touched upon, in various places throughout the book. No doubt with more time and space, all of them would have been more adequately discussed.

The author is to be especially congratulated on his steady refusal to be led astray by the recent emphasis on the more subjective aspects of the project; and on his consistent avoidance in his illustrations and discussions of the sentimentalities in which too much of our recent literature abounds. He is not seeking a method which mysteriously develops certain subtle and general qualities. He seeks a method through which plain objectives in the form of special abilities of known value may be most economically reached. He is undoubtedly right in his statement that "the provision for the natural setting of the teaching situation is the distinct contribution of the project method." In fact, the elaboration of this idea is the chief contribution of this book.

ERNEST HORN

State University of Iowa

SNEDDEN, DAVID. *Sociological determination of objectives in education.* Philadelphia: J. B. Lippincott Company, 1921. 322 pp.

In this book Doctor Snedden takes the position that the objectives of education must be determined, in the main, by data and methods that fall within the field of sociology. In order to equip the youth for the business of living, it is necessary to furnish training for the various occupations of adult life. Such training must be directed by studies of a sociological sort. Moreover, sociology can inform us as to the value or function of educational subject matter which has no direct bearing on vocation. "By the application of suitable sociological methods," for example, "it is entirely practicable to discover the scope and character of mathematical knowledge now used in any given 'standard of living' class or group, and, on the basis of the facts thus found and evaluated, to propose necessary or desirable improvements in processes of instruction and training to be applied to the rising generation" (p. 33).

In Doctor Snedden's exposition, this doctrine is linked up with the distinction between 'values of production' and 'values of consumption.' The former values have to do with vocation, and hence the standard is set by the needs and demands of society. The latter, on the other hand, "center definitely in cultivation of specific personal, intellectual, and aesthetic interests—the resources wherewith we enrich our leisure time, our individual lives. . . . They should at least establish abiding cultural interests—appreciations, tastes, enthusiasms, even hobbies—in literature, science, foreign languages, and history" (p. 83). This distinction in values points to a corresponding distinction in educational subjects or courses. One group, labelled "A class" subjects, "should lead very directly to processes and capacities known to be of use to the individual in adult life, or to the society of which he shall be a part" (p. 70). The second group, called "B class" subjects, "are those primarily which we follow, and to the degrees only which we follow them, because of innate and easily stimulated desire" (p. 71). The subjects in the former group are "hard"; they are properly imposed only because of obvious usefulness, and they are characterized by rigid method and by drills and tasks which are aimed at doing, at facility of expression in action. The subjects of the latter group, on the contrary, are directed rather towards absorption, or assimilation (p. 49), and so they partake more of the nature of "high-grade play."

It is this distinction between values of production and values of consumption that constitutes the basic principle of Doctor Snedden's philosophy of education. From this distinction he deduces the classification of subject matter into "A class" subjects and "B class" subjects, and also the further contrast between hard work or drill and "high-grade play." The former aims at "expression in action," while the latter is directed towards "assimilation" or "absorption." It is rather curious to note that in this exposition practically no account is taken of that considerable body of recent psychology which converges upon the conclusion that all knowing is doing, and which, accordingly, would repudiate the distinction in question. The trend of Doctor Snedden's doctrine at this point is not forward, but backward, toward the Lockean conception of the mind as just a wax tablet or sheet of white paper, upon which the environment may write what it will. Perhaps his position is defensible, but the psychology which underlies it has been too much under fire of late years to be taken for granted without argument.

But this is not the only point at which Dr. Snedden leans heavily upon a dubious tradition. His distinction between vocational and non-vocational or cultural subjects

is another instance of the same thing. The distinction is plausible only as long as we assume the validity of the historic opposition between vocation and culture. In spirit or emphasis Doctor Snedden's doctrine does not differ materially from the familiar notion that business or vocation is a thing quite apart from the nurture of the soul. The distinction presumably has a certain validity if we think of culture as, in the main, a private or subjective affair, in that it is directed towards the cultivation of detached, inner appreciations, which have no significant connections with the world of affairs. If we accept Dewey's definition of culture as "the capacity for constantly expanding in range and accuracy, one's perception of meanings," the opposition disappears and with it the hard and fast classification of subject matter into "A class" subjects and "B class" subjects. The efficacy of a plumber, for example, is determined in part, no doubt, by his training in the specific activities that constitute plumbing, but only in part. It is likewise determined by his conception of the relation that should obtain between employer and employee, by his standards of honesty, by his aesthetic appreciation, in short by his ability to see the meaning of his vocation thru its relations with a wide social context. To classify subjects offhand as vocational or non-vocational is to encourage the spirit of narrow vocationalism and to give aid and comfort to the exponents of an equally narrow and abstract "culture," against which Doctor Snedden so emphatically registers his opposition.

One further point should be noted. For the determination of educational objectives, the subject of sociology is, to Doctor Snedden, as the cloud by day and the pillar of fire by night. As far as vocational subjects are concerned, this view is plausible enough, provided that we take the meaning of vocation in a sufficiently narrow sense. Taken in this way, we can ascertain what is required for a given vocation as objectively and impersonally as we can determine the constitution of water or calculate the distance from the earth to the moon. But if we conceive of vocational training as something more than fitting a man to a groove, so that he can go through life with a minimum of intelligence, the matter takes on a different aspect. Doctor Snedden himself implies this when he says, in connection with non-vocational subjects, that the curriculum must be constructed "on the basis of the facts thus found *and evaluated*," but the significance of the evaluating evidently escapes his attention. He seems to take for granted that the evaluation of a more or less incidental task, which sociology is entirely willing to take off our hands and thus relieve us of all responsibility in the matter.

Given the same sociological method, there would presumably be no difference in the objectives of such men as Bismarck and Lincoln. Doctor Snedden gives no hint that the determination of objectives differs in any important respect from the process of prospecting for oil or coal. To be sure, the objectives are not supposed to be standing around in obscure places, waiting to be discovered, but there are facts to be discovered, and these facts, when brought to light, decide the question of objectives. Perhaps this too may be conceded. But can we assume that different individuals will discover the same method? Are we, or are we not, obliged to make allowance for a personal factor which somehow falls outside the field in which the sciences ply their trade, and which, because it tips the balance in favor of accepting certain facts as facts, has the last word in the determination of objectives?

This does not mean, of course, that the selection of objectives is to be left to random choice or caprice. But it does mean that the adoption of a gospel of brotherly love, of caring for the weak and of subordinating personal interests to group interests, or the adoption of any other gospel, involves certain factors which inevitably introduce

individual differences. Different men are bound to attach different meanings to the experience of freedom and responsibility, and to the desirability of cultivating an attitude of broad tolerance and sympathetic understanding, which means that they are bound to report differently upon the facts of experience. No scientific method, however rigorous, can provide an escape from these individual differences. The only rational check that can be applied lies in the organization of the facts which we happen to recognize into a coherent system, which then constitutes our philosophy of life and which, on occasion, becomes the basis for our determination of the objectives of education. It is simply self-deception to assume that in the determination of objectives the investigator can keep his deepest, nonscientific self out of the situation. To make such an assumption is scientific method gone mad.

It is easy to understand the present-day reaction against philosophy in its relation to education. In the past philosophy has served only too often as a substitute for the painstaking accumulation of fact by which educational theory and practice must be informed and directed. Perhaps Doctor Snedden is right when he says that "the philosopher can not, of course, be very patient of these attempts thus to consider democracy analytically." (p. 290). If I may trust my own observation, however, he is less disposed to be irritated by analysis than by the lack of it. The classification of values proposed by Doctor Snedden, the contrast between expression and absorption, and the distinction between work and play, are all sadly lacking in serious analysis. They are all conceptions that have had a long and evil history, and their weaknesses have been pointed out on many occasions and by many critics. Instead of substituting a more profound analysis, Doctor Snedden revamps these old conceptions and offers us a questionable philosophy of education, thinly disguised as sociology. I am glad of the opportunity to express my sincere admiration for the clearness and suggestiveness of Dr. Snedden's exposition; but I find myself wholly unable to agree with him as to the validity of his method and of his conclusions.

B. H. Bode

Ohio State University

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors.

Something from Minneapolis

According to Assistant Superintendent W. F. Webster, Minneapolis is about to establish a Bureau of Research. It is evident, however, that without a bureau organization Minneapolis has already been doing some of the work which a bureau ordinarily undertakes. Readers will be interested, for example, in Superintendent Webster's article in the current number of "The League Scrip." It is entitled "A Statistical Story." The author has at his disposal the records of pupils admitted to each of the various grades

for the past twenty years. His opportunity to study elimination is therefore unusual. It is not often true that reliable educational data are available for historical treatment. When, as in this case, the data include not only enrollments but numbers retarded, numbers promoted, and facts as to cost, the opportunity for significant reporting is obvious.

Research in Teacher Training

President H. A. Brown of the Oshkosh Normal School is evidently concerned because the light of investigation is not being turned on problems of teacher training as often or as fully as it should be. We agree with him. Anything that can be done in the selection and training of teachers is done at the very focal point of our educational enterprise. It is certain that the training of teachers is a topic on which no one can be dogmatic. The very diversity of the types and amounts of training indicates the lack of agreement. There must be best ways in which teachers may be trained. The field of teacher training is a promising one and we feel sure that any research worker who enters it will feel amply repaid for having done so.

Research in Higher Education

We are receiving a number of letters and reports which indicate that educational research is being applied to higher educational institutions. There is every reason why this should be so. The problems are scarcely less numerous that they are with reference to elementary and secondary education. The treatment of these problems, however, has hitherto been largely characterized by prejudice and an absence of objectivity. It seems to be especially true that the Bureaus of Educational Research which have been established in universities are being called upon to turn their attention to the condition of the institutions to which they belong. Perhaps research like charity may well begin at home.

Graduate Work and Salary Schedules

A recent correspondent raises the question of what is being done in the public schools to encourage teachers to pursue advanced work subsequent to graduation from a college or university. In other words, the question is whether it is possible for a teacher who has received training additional to that required for the baccalaureate degree to receive a higher salary or more rapid promotion, than the teacher who has not had this additional training. We shall be glad to receive information from cities whose salary schedules or whose regulations concerning promotion take into account the question of graduate work. It is our belief that account should be taken of it.

The Department at Okmulgee

Miss Velda Bamesberger, Director of the Department of Educational Statistics at Okmulgee, Oklahoma, has an organization which seems well adapted to serve the schools of a moderate sized city. Her department consists of three full-time people—a psychologist, a general assistant, and herself. It is surprising how much a well-organized department of this size can do if it succeeds in enlisting the support of the teachers—as Miss Bamesberger has apparently done. She says in a recent letter: "I have been devoting a great deal of my time to devising records, etc., for the system, and have also completely classified the seventh and eighth grades on the basis of intelligence tests, educational tests, and school marks. Our recent educational tests show that the classification is working well, and the high-school principal and teachers are more than pleased with it."

A Professional Duty

In the November number of the Detroit Educational Bulletin and under the above caption, Superintendent Frank Cody of the Detroit public schools presents vigorous plea to his teachers to join the National Education Association. After reviewing clearly and concisely the change that has taken place in the last four years in the public attitude toward education in general and toward teachers in particular, he shows that this change was largely brought about through the work of the National Association. He further points out that only if every teacher is solidly behind the N.E.A. can that Association furnish the leadership necessary to secure adequate public support for the schools of today and for the still better schools of tomorrow.

The article is illustrated (if this term is properly used) by the following quotation from Andrew Carnegie, which is printed in bold face, old English type: "You cannot push a person up a ladder unless he is willing to climb a little himself."

Bucyrus Survey Bulletin

Superintendent J. R. Patterson of Bucyrus, Ohio, has sent us one of the most pretentious survey bulletins which we have received. It is number one of volume three of his series. Superintendent Patterson has been conducting a cycle of three surveys—fall, mid-year, and spring, during the past two years; and the present bulletin gives data on the contents of the program. The bulletin is designed for the information of his teaching and supervisory staff and to disseminate survey notions among local citizens and interested fellow workers elsewhere. The contents reveal explanatory notes on statistics sufficient to enable the teacher, untrained in this line, to understand the meaning of the terms used in the body of the discussion.

The report of children's abilities involves column spelling from the Ayres-Buckingham scale, Courtis, Woody-McCall, and Monroe arithmetic, English Composition, Monroe reading, and hand-writing scored by the Ayres scale. There is also a section on age-grade conditions. It is unfortunate that, like so many others, Superintendent Patterson confuses retardation with overageness and acceleration with underageness, though his final statement to his last chart reveals the fact that he recognizes the real meaning of an age-grade table.

We hope that Superintendent Patterson will favor us with a copy of his survey bulletin, Vol. III, No. 2 when he completes it. We shall look forward with interest to the results contained therein and the treatment of the data which he will make available.

Another Survey Bulletin

There has also come to us a Survey Bulletin from the Kent Ohio city schools reporting the use of standardized tests during the second semester of the past year. The testing program here was somewhat more ambitious than the one at Bucyrus, since general intelligence tests as well as achievement tests were used, and the measurement of achievement involved a larger number of subjects. The subjects in which measurement was made were spelling, writing, arithmetic, composition, silent reading, geography, history, music, first-year algebra, second-year geometry, and third-year physics.

The bulletin, like the one previously described, includes numerous tables and graphs together with pointed comments on the scores and on the remedial measure which will be undertaken.

We have just received from W. H. Pillsbury of the Buffalo (N. Y.) city schools, a mimeographed copy of the "Buffalo Teachers' Library." This not only lists what

The Buffalo Teachers' Library

was thought by the committee of Buffalo teachers and administrators to be the best books and magazines for the teachers but also gives a brief review of each book. A sentence or two which shows the main problem and method of attack makes the list more valuable than the name of the book alone. The material has been classified under the following heads: (a) general reading; (b) elementary subjects; (c) secondary subjects; (d) vocational schools; (e) intermediate schools; (f) extension work; (g) teachers' organizations; (h) magazines for teachers; and (i) book companies. The first four of these large divisions is further subdivided and the material appropriately classified.

The work of compiling this library was done some two years ago, but lack of funds prevented it being printed at that time. It was mimeographed in order that it might be placed in the hands of the teachers. In his letter to us Mr. Pillsbury made the following statement: "We propose to revise it from time to time and are putting it out now with the hope that we will get suggestions and changes which will make each edition a little improvement on the preceding one. Any suggestions which may occur to you in the way of eliminations, additions, substitutions, or changes of method of preparation of the list, will certainly be most gratefully received."

The list, as it stands, is valuable to any of our readers. We suggest that superintendents of schools and research directors write Mr. Pillsbury for a copy, and having received it, show their appreciation by sending to him such suggestions as may occur to them.

English in Secondary Schools

A Report on the Conditions of the Teaching of English in the Secondary Schools of New Jersey (published by the New Jersey Association of Teachers of English and obtainable from the Secretary, Miss Mabel A. Tuttle, Linden, N. J.) will be of interest to all those concerned with high-school administration for two reasons: First, it presents in small compass a rather complete survey of the English situation in a state, treating details of schedule, curriculum, methods, etc. Second, it presents most frankly the attitude of the English teachers themselves on these questions.

In addition, the form of the report lends itself to ready comprehension. A short summary gives a bird's eye view of the situation, then each point is explained by numerous representative quotations from the questionnaires on which the report is based, and finally a schedule of eighteen "Recommendations" gives very definitely the standards and conditions which the teachers as a body think should prevail in high-school English teaching. With so much dissatisfaction abroad in regard to the results of English work, all those concerned should take this opportunity of securing further insight into existing difficulties and of learning what remedies an active body of teachers have tried.

E. W. Dolch, Jr.

University of Illinois

A Different Kind of Report Card

Superintendent E. B. Sellew of Middletown, Connecticut, has furnished us with a copy of his special report to the Board of Education and to parents of pupils in the Middletown schools. Since it is a radical departure from the usual form of report card or leaflet, it was thought that our readers will be interested in it.

The report is a four-page folder. The first page contains a black space for the name of the parent to whom the report is being sent and a general explanation of the

importance of school work with a plea for cooperation. Page two lists the subjects taught in high school together with the total number in each class and the number making the various grades or marks. For example, English I B, which means the first semester of first-year English, had 248 members, 47 of whom received grades of 90 to 100, 49 received 85 to 89, 105 received 75 to 84, 37 received 70 to 74, 10 received 60 to 69, and none received below 60%.

Grades of 85 to 100 are designed as honorary grades while any grade below 70 is a failing grade. The group in which the child is located is indicated by a red line. This calls the attention of the parent vividly to the number of students in the same grade with his child and to the number who received marks higher or lower than his child received.

Pages three and four give additional information concerning the schools, page four showing the school census by grades and the residence of high-school pupils with percent who are paying tuition.

While the report may be somewhat difficult for many of the parents to read, it is worth considering because of the large amount of information which is thus placed in the hands of patrons of the school.

A Few Data on the Use of the Stanford Revision of the Binet-Simon Tests by Halves

Apropos of the article in the September issue of this journal by Messrs. Otis and Knollin, the following data are submitted upon some of the same points covered by the article. In the course of several years of testing the writer accumulated a number of complete records of tests of pupils and became interested in making a comparison between the pupils' mental ages computed from the first and second halves of the tests. That is, each pupil's test record was divided into two parts, one representing his performance upon the first three or four, as the case may be, of the tests for each mental age level and the other his mental age as based upon the last three or four tests for each age. In doing this each test was, of course, given weight for just twice the number of months given when all the tests are used together. At the time of testing the pupils whose records are here treated, the writer had no thought of making such a study of their records and hence the giving of the tests could have been in no way affected by such a plan.

The records of 182 pupils were complete so that they could be included in this study. These pupils were well distributed through all the eight elementary grades and had mental ages ranging from four years and eight months to seventeen years and two months. After dividing their records into first and last halves by years as described above, the ages by the first half range from four years to sixteen years two months and by the second half from five years four months to eighteen years two months. The median age by the second half is 2.8 months higher than that by the first half. The average is 4.8 months higher. This latter figure seems to be unduly affected by a few rather freakish cases most of which occur at the mental ages of four and five and eleven and twelve. Except for these there is no apparent tendency for the age level of the pupils to make very much difference in their comparative mental ages by the first and second halves of the tests.

The following table gives the total distribution of differences in ages between the two halves. Positive differences mean that the ages by the second half are greater.

DIFFERENCES BETWEEN MENTAL AGES BY THE FIRST AND SECOND HALVES OF THE TESTS

Months	Number of Cases
−24–21	1
−20–17	1
−16–13	0
−12– 9	5
− 8– 5	17
− 4– 1	25
0	23
+ 1– 4	33
+ 5– 8	26
+ 9–12	22
+13–16	10
+17–20	8
+21–24	4
+25–28	1
+29–32	2
+33–36	2
+37–40	0
+41–44	1
+45–48	1
Total	182

As is stated above the median difference when signs are taken into consideration is 2.8 months. Neglecting signs and merely considering the absolute value of the differences, the median difference is 5.5 months. The average absolute difference is 8.2 months, being made much larger than the median by a few rather extreme cases. It will be noted that the seven greatest differences are all differences by which the second half gives a greater age than the first half. Also these seven are scattered amongst six different age levels, so it is apparent that the tests at any one age are not at fault here. The unusual results from the tests at ages four and five and eleven and twelve which were referred to above were not due to a few such extreme differences but to a larger number of differences of from eight to sixteen months.

The product-moment coefficient of correlation between the ages by the two halves is .92 with a P. E. of .01. The probable error of measurement[1] is found to be 5.5 months. This agrees very closely with the median absolute difference of 5.5 months mentioned above, and also is in substantial agreement with the results found by Otis and Knollin. This probable error of measurement is between four and five percent of the average mental age. This may be compared with seven percent which is the smallest probable error of measurement of any group intelligence test that the writer has noted.

C. W. Odell

University of Illinois

[1] The Probable Error of Measurement is obtained by the formula $P.E.M. = .6745\sigma\sqrt{1-r}$. For σ the average of the two standard deviations is used. r is the coefficient of reliability, that is, the coefficient of correlation between the first half and second half.

National Association of Directors of Educational Research

(E. J. Ashbaugh, *Secretary and Editor*)

Grand Rapids, Michigan.—Charles D. Dawson, Assistant Superintendent in charge of Research, has sent us a study of the marks given high-school pupils for the second semester of the present year. Grand Rapids has just changed its marking system from a percent basis to a five-letter basis. A rather extensive argument is included showing that the whole system of marking passing pupils between 75% and 100% is unscientific because the best and the poorest pupils often differ by more than 25% and because the different marks within these 25 units were not evenly distributed.

The new system gives the marks, A, B, C, D, and E and "Condition" with percentages of 10.5, 30.0, 30.0, 17.0, 9.5, and 3.0 respectively. This shows the marks to be skewed toward the upper end with the exception of the E value. Since the extent of this skewness increases from the 9th to 12th grades, it is held that the marks agree in practice with the theory that the successive grades in the high school represent greater and greater selection.

A great deal of variation is shown in the distribution of marks in the various subjects. It is stated, that "This is due mainly to variations in the type of students of which the different classes are made up, to the personal equation of the teachers who do the marking, and to the degree of difficulties of the work necessary to earn a passing mark in the different subjects." It is suggested that the teachers examine carefully the group of failures in their subjects as well as confer with one another concerning the use of the scale.

Those interested in making a study of their own distribution of marks, might well write Mr. Dawson for a copy of this report.

West Allis, Wisconsin.—T. L. Torgerson, Director of the Department of Educational Measurements, has devoted Bulletin No. II of his department to tests and measurements. He states that the bulletin is issued to give the new teachers a general idea of the scope and aims in the field of tests and measurements. A statement is made relative to general intelligence and some fundamental terms, definitions and underlying principles are set forth. The following statement is underscored for emphasis and is doubtless intended to be the key note in the measurement work for this year: "The pupils of the West Allis schools this year will be measured in terms of their own ability. The goal or standard for each pupil to reach is an achievement approximating 100% in efficiency."

His review of the work last year shows measurement in spelling, silent-reading, composition, arithmetic, and algebra. Statements are made regarding methods used, results obtained, and reaction of teachers to remedial work. Extensive intelligence testing was also done, and the following statement is made concerning the utilizing of results:

1. Mental age of the pupil used as a basis for grade placement.
2. Intelligence quotient of the pupil used as a basis for classification.

3. Intelligence and achievement compared to determine the achievement quotient of each pupil.

4. Aid for individual diagnosis.

5. Permanent record for future diagnostic purposes.

6. Discovery of sub-normal pupils.

A tentative outline of work for the year 1921-1922 includes for the high school a testing program in geography, history, and algebra; a testing program for the improvement of spelling, silent-reading, arithmetic, and handwriting; diagnosis of Clapp's Language test; a prognostic test of mathematical ability; a study of the vocabulary of the high-school pupils; testing for the appreciation of poetry in the junior high school; one group intelligence test and a study of the distribution of marks.

In the elementary schools a program of testing for remedial work in spelling, silent-reading, arithmetic, handwriting, language, and oral reading will be conducted. There will also be intelligence testing, a study of school marks, and a continuation of the study of the promotion of pupils on the basis of mental age. Closer supervision of the special help period for individual instruction will be given.

Denver.—Miss Emma M. Brown, Director of the Department of Measurements and Standards, has recently sent us two bulletins. Bulletin No. 2 deals with daily programs. Summarized, the general conclusions reached after a study of the teachers' daily programs are as follows:

1. There is a lack of uniformity as to accuracy, distribution of subjects, free time, and number of pupils.

2. Due to the fact that programs are not made accurately and carefully, there is much confusion as to interpretation.

3. A study of the principles involved would be helpful alike to principals and teachers.

4. Supervision of program making by the principal tends effectively to unify the work of the school in the following ways: (a) distribution of teachers' load as to teaching time, number of pupils, and free time; (b) distribution of pupils' time among various subjects; (c) reduction in amount of unsupervised study; (d) provision for coaching pupils who have been absent, or for helping backward or accelerated pupils; (e) variation in course of study to meet local needs; (f) provision for mental and educational testing and guidance; and (g) a more scientific placement of subjects. The tabular material presents data from which these conclusions were drawn. The whole bulletin is suggestive to those who wish to improve the daily programs of teachers in their school systems.

Bulletin No. 3 reports two tests in spelling in elementary and junior high schools, grades II-A to VIII-A inclusive, September 1920 and January 1921. The bulletin not only presents in figures and graphs the conditions in the various Denver Schools, and in the city as a whole, but also includes deductions, suggestions for improvement, and some general principles for the benefit of the teachers.

The following are some of the results:

(1) That girls spelled better than boys; (2) that there was a difference between the ability in content and list spelling which gradually decreased in the higher grades; (3) that locality of school and size of class had little effect; (4) that schools with foreign populations were handicapped in the lower grades, but overcame the handicap in the upper grades.

Under suggestions for improvement, the following statements, while not new, will bear repetition:

1. Secure automatic accuracy for words of greatest frequency in school work, social usage, and adult correspondence.

2. Give preliminary tests in order that time may not be wasted in teaching children words which they already know.

3. It is more important that a pupil should have a definite and usable method of learning than that a high technique should be developed on the part of the teacher.

4. "Trouble spots" occur in some words and not in others and frequently change in the same word from one grade to another. For this reason it is doubtful if it is worth while to spend much time in calling attention to them.

5. Homonyms may be effectively taught separately in initial presentation and together when words have been confused.

Kansas City, Missouri.—George Melcher, Assistant Superintendent in charge of Research and Efficiency, has favored us with two of his circular letters to principals and teachers and with some data on penmanship work which the Bureau has been doing. Because these communications contain suggestions which may be of service to the other members, they are described in this department.

Members of the Association will probably remember that the Kansas City Bureau some time ago arranged a scale for the measurement of handwriting. The work was done by the same general method that Thorndike used, and the resulting scale is printed in the same form. It includes, however, grade standards and suggestions on the handwriting scale, which make it particularly valuable to the principals and teachers of the local schools. The Bureau now issues a set of fifty specimens of handwriting which have been graded by 89 judges. A median value, as determined by these ratings, has been assigned to each specimen. The specimens are designed to help teachers, principals, and supervisors in standardizing their own scoring of quality of handwriting.

While smaller systems may not wish to undertake as comprehensive a scheme of training for their teachers, yet even the smallest might well consider the advisability of accumulating a set of samples well distributed in quality and scored often enough to determine well established values. These would assist materially in the training of teachers in the use of the scale.

One of the two bulletins mentioned above presents date on overageness. The following points seem worthy of note: In 1896 the percent of overageness in the Kansas City schools was 57, while in 1921 it was 27.5. Since in 1914 the percent of overageness was 45, the rate of reduction in the last seven years has been more than three times as rapid as in the preceding eighteen years. In 1896, 35.6 percent of the children who entered the schools were graduating from the elementary grades; in 1914, 55.2 percent and in 1921, 81.8 percent. The question is raised whether this reduction in overageness will not lower the quality of school work. The answer is made that standardized achievement tests during the past few years have shown steady advancement while the reduction in overageness has been going on.

Many of our members will doubtless be interested in making a similar comparison in their own schools. Your secretary will appreciate it greatly if the findings are forwarded to him for use in this department.

The other bulletin deals with reports on the Courtis Arithmetic Tests. The report shows that the city-wide medians were practically the same in May 1921 as in May 1920. Of all the classes of the city 42.5 percent scored below normal. The following goal is set: "Every class should aim at the standard. Slow classes should reach the lowest normal score in each grade, average classes should reach the standard, and strong classes may score somewhat above standard. However, it is advised that classes should not exceed the upper normal scores in any grade." The upper and lower normal scores referred to are upper and lower quartiles. The author also calls the attention of all teachers and principals to the fact that one school which had always been in the lowest quartile until last year had brought all grades to normal or above. "This splendid piece of work was done in one year under only fair working conditions and with average pupils."

Cannot our readers furnish us with similar records?

University of Michigan.—Last spring, the University of Michigan, Bureau of Tests and Measurements, sent out a bulletin of inquiry to the school people of the state asking for guidance in regard to testing programs. The tabulation of the replies gives an interesting index of the preference which school people in that state are exhibiting in this field. The tabulation follows:

Of 47 replies, 46 wish to include an intelligence test.
" 42 " 25 wish to include at least Grades III to VIII in the program; 32 wish to include six or more of the school grades.
" 31 " 29 wish to include three or more school subjects, and 24 wish to include four or more, though sometimes these four include the intelligence test.
" 31 " 22 wish to use the Courtis Arithmetic Test.
" 33 " 19 wish to use the Monroe Silent Reading Test.
" 47 " 24 wish to use the Ayres Spelling Scale.
" 29 " 19 wish to use the Ayres Handwriting Scale.
" 27 " 19 wish to use the National Intelligence Test.
" 47 " 29 are willing to leave it to the Bureau to decide the order in which the tests shall be given.
" 47 " 44 agree to forward results to the Bureau.

On the basis of this information, the Bureau arranged a program for work in grades three to eight which included courtis Supervisory Arithmetic Tests, Monroe Silent Reading Tests, Buckingham Extension of the Ayres Spelling Scale and the National Intelligence Tests.

We shall look forward with interest to the report on this program and we shall welcome reports of similar programs from other Directors.

We take pleasure in announcing

THE TWENTY-FIRST YEARBOOK

of the

National Society for the Study of Education

¶The Twenty-First Yearbook of the National Society for the Study of Education will deal with the subject of intelligence tests, discussing their nature and history, the general principles underlying their use, and the administrative use of intelligence tests from the kindergarten through the university.

¶Among the authors of the Twenty-First Yearbook are:
E. L. Thorndike, Teachers College, Columbia University.
S. S. Colvin, Brown University.
H. W. Holmes, Graduate School of Education, Harvard University.
G. M. Whipple, University of Michigan.
Agnes L. Rogers, Goucher College.
Warren K. Layton, Public Schools, Detroit.

¶The standing of the authors will make the Twenty-First Yearbook one of the authoritative publications of 1922 on the subject of intelligence testing.

¶Though the Twenty-First Yearbook is satisfyingly thorough in its complete discussion of its subject, it is written in a clear style that will win favor with the thousands of teachers who are desirous of learning as much as possible concerning the testing of intelligence.

¶The Twenty-First Yearbook will be displayed for the first time at Chicago, February 25th, 1922, in the Gold Room of the Congress Hotel. Part II of the Twenty-First Yearbook will be discussed Tuesday evening, February 28th, in the Auditorium Theatre at a joint meeting of the National Society for the Study of Education and the Department of Superintendence of the N. E. A.

Advance orders should be sent to the

PUBLIC SCHOOL PUBLISHING COMPANY

BLOOMINGTON, ILLINOIS

THREE NOTABLE BOOKS FOR TEACHERS' READING AND STUDY

The Psychology of Arithmetic

By EDWARD L. THORNDIKE

Professor of Educational Psychology, Teachers College, Columbia University

Recognizing that the ordinary view of the arithmetical learning is obscure or inadequate as to specify knowledge desired, terminology of instruction, and ability to apply what is learned in arithmetic to the quantitative problems of life, Dr. Thorndike has made a brilliant analysis of the subject of arithmetical teaching and study to show what has been done erroneously or inadequately, and what should be done in reform.

Fundamentals of Education

By BOYD H. BODE

Professor of the Philosophy of Education, Ohio State University

Edited by William C. Bagley for THE MODERN TEACHERS' SERIES

What should be the aims and ideals back of a teacher's effort? What should present day education attempt? What standards of value should determine the materials of the curriculum, the organization of the school, the methods of instruction, the intimate and influential relations of the teacher with his pupils? A study of the philosophy of education will not answer all these questions but it should do much to help the teacher to answer them for himself. Professor Bode has succeeded admirably in making this quality the dominating feature of this book.

A Guide to the Teaching of Spelling

By HUGH CLARK PRYOR

Northern Normal and Industrial School, Aberdeen, South Dakota

and

MARVIN SUMMERS PITTMAN

Michigan State Normal College

This book is an attempt to evaluate and make available for busy teachers the results of recent investigations and studies in the subject of spelling. It indicates clearly the possibilities and the difficulties of improvement in spelling, outlines the mode of procedure, and suggests the proper motivation.

Coming Soon

Book: THE INTELLIGENCE OF HIGH SCHOOL SENIORS
Boraas: TEACHING TO THINK
Ingraham: THE COMMUNITY

Send for the Complete Educational Catalogue

The Macmillan Company

New York Atlanta San Francisco
Boston Chicago Dallas